

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2010-523979

(P2010-523979A)

(43) 公表日 平成22年7月15日(2010.7.15)

(51) Int.Cl.	F I	テーマコード (参考)
GO 1 N 33/574 (2006.01)	GO 1 N 33/574	D
GO 1 N 33/533 (2006.01)	GO 1 N 33/533	

審査請求 未請求 予備審査請求 未請求 (全 127 頁)

(21) 出願番号	特願2010-502159 (P2010-502159)	(71) 出願人	509274968
(86) (22) 出願日	平成20年4月7日 (2008.4.7)		オーレオン ラボラトリーズ, インコーポレイテッド
(85) 翻訳文提出日	平成21年10月2日 (2009.10.2)		アメリカ合衆国 ニューヨーク 1070
(86) 国際出願番号	PCT/US2008/004523		1, ヨンカーズ, ウェルズ アベニュー
(87) 国際公開番号	W02008/124138		28
(87) 国際公開日	平成20年10月16日 (2008.10.16)	(74) 代理人	100078282
(31) 優先権主張番号	60/922, 163		弁理士 山本 秀策
(32) 優先日	平成19年4月5日 (2007.4.5)	(74) 代理人	100062409
(33) 優先権主張国	米国 (US)		弁理士 安村 高明
(31) 優先権主張番号	60/922, 149	(74) 代理人	100113413
(32) 優先日	平成19年4月5日 (2007.4.5)		弁理士 森下 夏樹
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/923, 447		
(32) 優先日	平成19年4月13日 (2007.4.13)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 医学的状態の処置、診断および予測のためのシステムおよび方法

(57) 【要約】

医学的状態（例えば、癌）の出現（例えば、再発）の予測のための予測モデルにおいて臨床情報、分子情報、およびコンピュータ生成された形態情報（computer-generated morphometric information）を使用する方法およびシステムを提供する。1つの実施形態では、フィーチャ（1つまたは複数（例えば、全て）の生検グリーソンスコア、精嚢浸潤、嚢外拡大、術前PSA、優位前立腺摘除グリーソン分類（dominant prostatectomy Gleason grade）、AR+上皮核の相対領域、上皮核の形態計測、および上皮細胞質の形態計測が含まれる）に基づいた前立腺癌再発を予測するモデルを提供する。

FIG. 17

Features Selected for Prediction of PSA Recurrence (Example 8)

Clinical	Biopsy Gleason Score	-25.60
	Seminal Vesicle Invasion	-22.69
	Extracapsular Extension	-4.37
	Preoperative PSA	-13.81
	Dominant Prostatectomy Gleason Grade	-10.93
Molecular	Relative Area of AR+ Epithelial Nuclei	-10.64
Morphometric	Texture of Tumor Epithelial Nuclei	-19.21
	Texture of Tumor Epithelial Cytoplasm	18.19

【特許請求の範囲】

【請求項 1】

患者の前立腺癌の再発リスクを評価するための装置であって、前記装置が、
患者のデータセットを評価し、それによって前記患者の前立腺癌の再発リスクを評価するよう構成された前立腺癌再発予想モデルを含み、前記モデルが、以下のフィーチャの群：

- 生検グリーソンスコア；
- 精嚢浸潤；
- 嚢外拡大；
- 術前 P S A；
- 優位前立腺摘除グリーソン分類；
- A R + 上皮核の相対領域；
- 組織画像由来の上皮核の形態計測；および
- 組織画像由来の上皮細胞質の形態計測

から選択される 1 つまたは複数のフィーチャに基づく、装置。

【請求項 2】

前記上皮核の形態計測が腫瘍上皮核のテクスチャの測定を含む、請求項 1 に記載の装置。

【請求項 3】

前記上皮細胞質の形態計測が腫瘍上皮細胞質のテクスチャの測定を含む、請求項 1 に記載の装置。

【請求項 4】

前記モデルが前記群由来の 2 つのフィーチャに基づく、請求項 1 に記載の装置。

【請求項 5】

前記モデルが前記群由来の 3 つのフィーチャに基づく、請求項 1 に記載の装置。

【請求項 6】

前記モデルが前記群由来の 4 つのフィーチャに基づく、請求項 1 に記載の装置。

【請求項 7】

前記モデルが前記群由来の 5 つのフィーチャに基づく、請求項 1 に記載の装置。

【請求項 8】

前記モデルが前記群由来の 6 つのフィーチャに基づく、請求項 1 に記載の装置。

【請求項 9】

前記モデルが前記群由来の 7 つのフィーチャに基づく、請求項 1 に記載の装置。

【請求項 10】

前記モデルが前記群由来の全 8 つのフィーチャに基づく、請求項 1 に記載の装置。

【請求項 11】

前記モデルが前記患者の医学的状態の出現リスクを示す値が得られるよう構成されている、請求項 1 に記載の装置。

【請求項 12】

前記モデルが A R + 上皮核の相対領域のフィーチャに基づき、前記フィーチャが免疫蛍光を示す組織画像のコンピュータ分析に基づく、請求項 1 に記載の装置。

【請求項 13】

患者の前立腺癌の再発リスクを評価する方法であって、前記方法が、
前立腺癌再発を予測するモデルを使用して前記患者のデータセットを評価する工程であって、前記モデルが、以下のフィーチャの群：生検グリーソンスコア、精嚢浸潤、嚢外拡大、術前 P S A、優位前立腺摘除グリーソン分類、A R + 上皮核の相対領域、組織画像由来の上皮核の形態計測、および組織画像由来の上皮細胞質の形態計測から選択される 1 つまたは複数のフィーチャに基づき、それにより、前記患者の前立腺癌の再発リスクを評価する工程を含む、方法。

【請求項 14】

前記モデルが前記群由来の 2 つのフィーチャに基づく、請求項 1 に記載の装置。

前記上皮核の形態計測が腫瘍上皮核のテクスチャの測定を含む、請求項 1 3 に記載の方法。

【請求項 1 5】

前記上皮細胞質の形態計測が腫瘍上皮細胞質のテクスチャの測定を含む、請求項 1 3 に記載の方法。

【請求項 1 6】

前記患者の前立腺癌の再発リスクを示す値を出力する工程をさらに含む、請求項 1 3 に記載の方法。

【請求項 1 7】

免疫蛍光を示す組織画像のコンピュータ分析に基づいて A R + 上皮核の相対領域のフィーチャを生成する工程をさらに含む、請求項 1 3 に記載の方法。

10

【請求項 1 8】

コンピュータ読み取り可能な媒体であって、

前立腺癌再発を予測するモデルを使用して患者のデータセットを評価し、それにより、前記患者の前立腺癌の再発リスクを評価する工程であって、前記モデルが、以下のフィーチャの群：生検グリーンスコア、精嚢浸潤、嚢外拡大、術前 P S A、優位前立腺摘除グリーンソン分類、A R + 上皮核の相対領域、組織画像由来の上皮核の形態計測、および組織画像由来の上皮細胞質の形態計測から選択される 1 つまたは複数のフィーチャに基づく、評価工程

を含む方法を実行するための前記コンピュータ読み取り可能な媒体上に記録したコンピュータ実行可能命令を含むコンピュータ読み取り可能な媒体。

20

【請求項 1 9】

根治的前立腺摘除を受けた後の患者の臨床的な失敗リスクを評価するための装置であって、前記装置は、以下：

患者のデータセットを評価し、それによって前記患者の臨床的な失敗リスクを評価するように構成された臨床的な失敗を予想するモデルを含み、前記モデルが、以下のフィーチャの群：

優位前立腺摘除グリーンソン分類；

リンパ節の浸潤状態；

組織画像由来の管腔の形態測定；

組織画像由来の細胞質の形態計測；および

A R + / A M A C R - 上皮核中の A R の平均強度

から選択される 1 つまたは複数のフィーチャに基づく、装置。

30

【請求項 2 0】

前記管腔の形態計測が平均管腔周囲長を含む、請求項 1 9 に記載の装置。

【請求項 2 1】

前記管腔の形態計測が相対管腔領域を含む、請求項 1 9 に記載の装置。

【請求項 2 2】

前記モデルが平均管腔周囲長および相対管腔領域を含む少なくとも 2 つの管腔の形態測定に基づく、請求項 1 9 に記載の装置。

40

【請求項 2 3】

前記モデルが A R + / A M A C R - 上皮核中の A R の平均強度のフィーチャに基づき、前記フィーチャが免疫蛍光を示す組織画像のコンピュータ分析に基づく、請求項 1 9 に記載の装置。

【請求項 2 4】

前記モデルが前記群由来の 2 つのフィーチャに基づく、請求項 1 9 に記載の装置。

【請求項 2 5】

前記モデルが前記群由来の 3 つのフィーチャに基づく、請求項 1 9 に記載の装置。

【請求項 2 6】

前記モデルが前記群由来の 4 つのフィーチャに基づく、請求項 1 9 に記載の装置。

50

【請求項 27】

前記モデルが前記群由来の5つのフィーチャに基づく、請求項19に記載の装置。

【請求項 28】

前記モデルが、患者の臨床的な失敗リスクを示す値を生成するように構成されている、請求項19に記載の装置。

【請求項 29】

根治的前立腺摘除を受けた後の患者の臨床的な失敗リスクを評価する方法であって、前記方法は、以下：

前立腺摘除後の臨床的失敗を予測するモデルを使用して前記患者のデータセットを評価する工程であって、前記モデルが、以下のフィーチャの群：優位前立腺摘除グリーンソン分類、リンパ節の浸潤状態、組織画像由来の管腔の形態測定、組織画像由来の細胞質の形態計測、およびAR+/AMACR-上皮核中のARの平均強度から選択される1つまたは複数のフィーチャに基づき、

それにより、患者の臨床的な失敗リスクを評価する工程を含む、方法。

【請求項 30】

前記管腔の形態計測が平均管腔周囲長を含む、請求項29に記載の方法。

【請求項 31】

前記管腔の形態計測が相対管腔領域を含む、請求項29に記載の方法。

【請求項 32】

前記モデルが平均管腔周囲長および相対管腔領域を含む少なくとも2つの管腔の形態測定に基づく、請求項29に記載の方法。

【請求項 33】

前記患者の臨床的な失敗リスクを示す値を出力する工程をさらに含む、請求項29に記載の方法。

【請求項 34】

免疫蛍光を示す組織画像のコンピュータ分析に基づいてAR+/AMACR-上皮核中のARの平均強度のフィーチャを生成する工程をさらに含む、請求項29に記載の方法。

【請求項 35】

コンピュータ読み取り可能な媒体であって、前立腺摘除後の臨床的失敗を予測するモデルを使用して患者のデータセットを評価し、それにより、前記患者の臨床的な失敗リスクを評価する工程であって、前記モデルが、以下のフィーチャの群：優位前立腺摘除グリーンソン分類、リンパ節の浸潤状態、組織画像由来の管腔の形態測定、組織画像由来の細胞質の形態計測、およびAR+/AMACR-上皮核中のARの平均強度から選択される1つまたは複数のフィーチャに基づく、評価工程

を含む方法を実行するために前記コンピュータ読み取り可能な媒体上に記録したコンピュータ実行可能命令を含むコンピュータ読み取り可能な媒体。

【発明の詳細な説明】

【技術分野】

【0001】

(関連出願)

本出願は、2006年10月13日に出願された米国特許出願第11/581,052号(これは2005年10月13日に出願された米国仮特許出願第60/726,809号からの優先権を主張し、2005年3月14日に出願された米国特許出願第11/080,360号の一部継続出願であり、米国特許出願第11/080,360号は2005年2月25日に出願された米国特許出願第11/967,066号(現在では2008年1月22日発行の米国特許第7,321,881号)の一部継続出願であり、米国特許出願第11/967,066号は2004年2月27日に出願された米国仮特許出願第60/548,322号および2004年6月4日に出願された米国特許第60/577,051号からの優先権を主張する)の一部継続出願であり、2003年11月17日に出願さ

10

20

30

40

50

れた米国仮特許出願第60/520,815号からの優先権を主張する2004年11月17日に出願された米国特許出願第10/991,897号の一部継続出願であり、2003年7月21日に出願された米国仮特許出願第10/624,233号(現在では2006年2月7日発行の米国特許第6,995,020号)の一部継続出願であり、2003年11月18日に出願された米国仮特許出願第60/520,939号からの優先権を主張する2004年11月17日に出願された米国仮特許出願第10/991,240号の一部継続出願であり、そして2004年3月12日に出願された米国仮特許出願第60/552,497号、2004年6月4日に出願された米国仮特許出願第60/577,051号、2004年8月11日に出願された米国仮特許出願第60/600,764号、2004年10月20日に出願された米国仮特許出願第60/620,514号、2005年1月18日に出願された米国仮特許出願第60/645,158号、2005年2月9日に出願された米国仮特許出願第60/651,779号からの優先権を主張する。本出願はまた、2007年4月5日に出願された米国仮特許出願第60/922,163号、2007年4月5日に出願された米国仮特許出願第60/922,149号、2007年4月13日に出願された米国仮特許出願第60/923,447号、および2008年1月9日に出願された米国仮特許出願第61/010,598号による優先権を主張する。

10

20

30

40

50

【0002】

(発明の分野)

本発明の実施形態は、医学的状態(例えば、疾患または治療に対する応答性もしくは非応答性)の出現を予測するための予測モデルにおける臨床情報、分子情報、およびコンピュータ生成された形態情報(computer-generated morphometric information)を使用する方法およびシステムに関する。例えば、1つの実施形態では、本発明は、前立腺癌の再発を治療、診断、および予測するために臨床情報、分子情報、および形態情報を使用する方法およびシステムを含む。

【背景技術】

【0003】

(背景)

医師は、例えば、患者が医学的状態を経験する可能性が高いかどうかおよびそれはいつかといった問題や、一旦患者が疾患と診断されるとどのようにして患者を治療すべきかといった問題にまで亘って、多数の医学的判断を行う必要がある。患者に対する一連の治療を適切に決定すれば、例えば、患者が生存および/または回復する機会が増大し得る。同様に、事象の出現の予想により、個体の事象の治療を有利に計画することが可能である。例えば、患者が疾患の出現(例えば、再発)を経験する可能性が高いかが予想されると、医師はこの患者に適切な一連の治療を推奨することが可能である。

【0004】

伝統的には、医師は、医学的状態の出現の治療、診断、および予測のために医師の専門知識および訓練に大いに依存する。例えば、病理学者は、前立腺癌の進行および侵襲レベルを評価するためにグリーソンスコアリングシステム(Gleason scoring system)を使用し、このシステムでは、医師によって把握される顕微鏡下での前立腺組織の外観に基づいて癌が分類される。未分化型がより多数である前立腺組織サンプルに、より高いグリーソンスコアが与えられる(1)。グリーソン分類は病理学者によって信頼できると広く見なされているが、グリーソン分類は主観的なスコアリングシステムである。特に、同一組織サンプルを観察する病理学者が異なれば、解釈も異なる。

【0005】

医学的診断において医師を補助するための従来のツールは、範囲および適用が制限される。例えば、患者が根治的前立腺摘除を受けた後に前立腺癌治療に関する医師の決定を補助するためのツールは、血清ベースのPSAスクリーニング試験および一般化ノモグラム(generalized nomogram)に制限される。Kattan et al.の特許文献1によって開発された1つの術後ノモグラムは、泌尿器科医によって広く

使用されており、根治的前立腺摘除治療を受けた患者の7年疾患再発率を予測可能である。このノモグラムは、生化学的不全（すなわち、PSAレベルの増加）の可能性についての情報しか得られず、臨床的失敗（死亡）を予測しない。さらに、このノモグラムは、患者の容態が7年以内に再発する可能性が高いかどうかのみを予測し、7年間のうち患者の容態が再発し得る時期を予想しない。このノモグラムで使用される予後変数には、処置前血清PSAレベル、グリーソンスコア、および病理学者による前立腺被膜浸潤、切除縁、精嚢浸潤、およびリンパ節の状態微視的評価が含まれる。処置の失敗は、疾患の再発の臨床的証拠が存在するか、血清PSAが上昇するか、または補助療法を開始する場合に記録される。しかし、これらのノモグラムにはいくつかの制限が存在する。最も重要な制限は、最良のこれらのノモグラムでさえ、完全に区別されるモデル（一致指数（concordance index）= 1.0）と区別できないモデル（一致指数 = 0.5）との間の中間よりもわずかに良好にしか実施されないという点である。さらに、予測がコイントスほどの精度に過ぎないので、中間領域のノモグラム予測（7年無増悪生存率30～70%）を有する患者の約30%の転帰が不明確である。

10

20

30

40

50

【0006】

コンピュータの処理能力が有意に増大した、コンピュータ実行画像処理および画像分析の技術が出現している。多数の応用面において、単一画像から大量の定量的連続値フィーチャを自動的に抽出することが現実的にできるようになっている。フィーチャXは、あるAがB未満の場合において、フィーチャの値セットがAとBとの間の全ての数xを含む場合に、連続値であるといわれる。癌画像分析システムは、細胞学的検体の撮影画像について開発されている（2）（3）。しかし、かかるシステムは、細胞のみを取り込んでいるので、組織レベルで観察可能な構造情報の全てを利用せず、まして臨床情報および分子情報がその情報と組み合わせられない。組織レベルで異なる病理学的要素の構造を分析するための癌画像分析システムは提供されていないが、組織レベルの構造は、しばしば、各細胞の外観よりも診断（例えば、グリーソン分析）で重要な役割を果たす。したがって、病理学者は、前立腺の形状およびサイズを分析するために手作業の技術を用いて、癌の病理学的悪性度を決定している（4）。組織画像が典型的には細胞画像より複雑であり、理解するために専門分野の幅広い知識が必要であるという事実が、従来癌画像分析システムの欠点をより深刻にしている。

【先行技術文献】

【特許文献】

【0007】

【特許文献1】米国特許第6,409,664号明細書

【発明の概要】

【発明が解決しようとする課題】

【0008】

上記を考慮して、予測力が改善された医学的状态、応答、および他の医学的現象の出現を治療、診断、および予測するためのシステムおよび方法を提供することが望ましいであろう。医学的状态の治療、診断、およびその出現の予測のために組織レベルの情報を使用するコンピュータ実行システムおよび方法を提供することも望ましいであろう。

【課題を解決するための手段】

【0009】

（発明の概要）

本発明の実施形態は、医学的状态の出現を予測するための自動化システムおよび方法を提供する。本明細書中で使用する場合、医学的状态の出現の予測には、例えば、患者が癌などの疾患の出現（例えば、再発）を経験するかどうかの予測および/またはその経験する時期の予測、患者が1つまたは複数の療法（例えば、新規の医薬品）に応答する可能性が高いかどうかの予測、および任意の他の適切な医学的状态の出現の予測が含まれ得る。本発明の実施形態による予測を、例えば、患者のための適切な一連の治療を選択し、および/または患者の医学的状态を診断するために、医師または他の個人によって使用するこ

とができる。

【0010】

本発明の1つの態様では、医学的状態の出現を予測するモデルを生成するためのシステムおよび方法を提供する。予測モデルの生成は、その転帰が少なくとも部分的に知られている患者のコホートについてのデータを使用してサポートベクトルマシン (SVM) または神経回路網を訓練するための分析ツールの使用を含み得る。1つの実施形態では、トレーニングデータには、臨床データ、分子データ、およびコンピュータ生成された形態計測データが含まれる。本明細書中で使用する場合、特定の型の「データ」(例えば、臨床データ、分子データ、または形態計測データ)には、その型の1つまたは複数のフィーチャが含まれ得る。さらに、形態計測データを、組織の電子(デジタル)画像に関連するかこれに由来する任意のコンピュータ生成されたデータ(組織またはその一部の構造特性(例えば、領域、長さ、幅、稠密度、および密度)、組織またはその一部のスペクトル特性(例えば、赤色、緑色、青色(RGB)色成分値(color channel values)、輝度、およびチャンネルヒストグラム(channel histogram))、組織画像および/または同定された組織成分(例えば、上皮内接触面、内腔輪郭のフラクタル次元)のフラクタル特性、ウェーブレット分解係数の統計的特性、ならびに/または他の画像データ変換に関するデータが含まれるが、これらに限定されない)が含まれると定義する。他の実施形態では、トレーニングデータには、コンピュータ生成された形態計測データのみまたは臨床データとコンピュータ生成された形態計測データとの組み合わせが含まれる。

10

20

【0011】

1つの実施形態では、間質、細胞質、上皮核、間質核、管腔、赤血球、組織人工物(tissue artifact)、組織バックグラウンド(tissue background)、またはその組み合わせに関連する1つまたは複数のコンピュータ生成された形態計測的フィーチャに基づいた予測モデルを生成するためのシステムおよび方法を提供し、コンピュータ生成された形態計測的フィーチャのみまたは表4に列挙した1つまたは複数の臨床的フィーチャおよび/または表6に列挙した1つまたは複数の分子的フィーチャとの組み合わせに基づいて予測モデルを生成することができる。例えば、1つまたは複数のフィーチャを、関連するモデルが医学的状態を予測する能力に及ぼすフィーチャの影響を決定する分析ツールに入力することができる。モデルの予測力を増加させるフィーチャを最終モデルに含めることができるのに対して、予測力を増加させない(または減少させる)フィーチャを、検討事項から除去することができる。予測モデル開発の根拠として、上記形態計測的フィーチャを単独または表4および/または表6に列挙した臨床的フィーチャおよび/または形態計測的フィーチャとそれぞれ組み合わせて使用することにより、医師、他の個人、および/または自動化処理装置(例えば、組織画像分析システム)のリソースを、転帰と相関する可能性がより高く、したがって、最終予測モデルで有用な患者のフィーチャについてのデータの取得に向けることができる。

30

【0012】

本発明の別の態様では、コンピュータ生成された形態計測的データのみまたは臨床データおよび/または分子データとの組み合わせに基づく、患者に於ける医学的状態の出現リスクを評価するために患者のデータセットを評価する予測モデルを提供する。例えば、予測モデルは、入力として患者のデータセットを受信し、患者が医学的状態に関連する1つまたは複数の転帰を経験する可能性を示す「スコア」を出力することができる。

40

【0013】

1つの特定の実施形態では、1つまたは複数(例えば、全て)の以下: 生検グリーンソンスコア、精嚢浸潤、嚢外拡大、術前PSA、優位前立腺摘除グリーンソン分類(dominant prostatectomy Gleason grade)、AR+上皮核の相対領域(例えば、免疫蛍光を示す組織画像のコンピュータ分析に基づいて生成)、上皮核の形態計測、および上皮細胞質の形態計測が含まれるフィーチャに基づく、患者の前立腺癌の再発リスクを予測する予測モデルを提供する。

50

【0014】

別の実施形態では、1つまたは複数の以下：精嚢の関与、切除縁の関与、リンパ節の状態、腫瘍のアンドロゲン受容体（AR）染色指数、組織画像由来の上皮核の形態計測（例えば、総組織領域で除した上皮核によって占められる領域）、および/または組織画像由来の間質の形態計測（例えば、総組織領域で除した間質によって占められる領域）が含まれるフィーチャに基づく、患者の前立腺癌の再発リスクを予測する予測モデルを提供する。いくつかの実施形態では、モデルは、生検グリーンスコアのフィーチャおよび/または組織画像に由来する間質内のテクスチャの測定にさらに基づき得る。モデルは、患者のデータセットを評価し、それにより、患者の前立腺癌の再発リスクを評価することができる。

10

【0015】

別の実施形態では、間質、細胞質、上皮核、間質核、管腔、赤血球、組織人工物、組織バックグラウンド、またはその組み合わせに関連する1つまたは複数のコンピュータ生成された形態計測的フィーチャに基づく、疾患の出現または再発を予測するための予測モデルを提供する。予測モデルは、コンピュータ生成された形態計測的フィーチャのみまたは表4に列挙した1つまたは複数の臨床的フィーチャおよび/または表6に列挙した1つまたは複数の分子的フィーチャとの組み合わせに基づくことができる。

【0016】

別の実施形態では、図6に示す1つまたは複数の臨床的フィーチャおよび/または分子的フィーチャならびに1つまたは複数の以下の病理学的オブジェクト：赤血球、上皮核、間質、管腔、細胞質、および組織バックグラウンドについての1つまたは複数の形態計測的フィーチャに基づく、前立腺癌の再発を予測するための予測モデルを提供する。

20

【0017】

さらに別の実施形態では、図9に示す1つまたは複数の臨床的フィーチャおよび/または分子的フィーチャならびに1つまたは複数の以下の病理学的オブジェクト：赤血球、上皮核、間質、管腔、および細胞質についての1つまたは複数の形態計測的フィーチャに基づく、前立腺癌の再発を予測するための予測モデルを提供する。

【0018】

別の実施形態では、図11に示す1つまたは複数の臨床的フィーチャおよび/または分子的フィーチャならびに1つまたは複数の以下の病理学的オブジェクト：赤血球、上皮核、および間質についての1つまたは複数の形態計測的フィーチャに基づく、前立腺癌の生存性を予測するための予測モデルを提供する。

30

【0019】

他の実施形態では、予測モデルにより、組織サンプルが正常または異常であるかどうか決定することができるか、あるいは、患者が前立腺摘除後の臨床的な失敗を経験する可能性が高いかどうかを予測することができる。1つの特定の実施形態では、1つまたは複数（例えば、全て、1つ、2つ、3つ、4つなど）の以下：優位前立腺摘除グリーンソン分類、リンパ節の浸潤状態、少なくとも1つの管腔の形態計測、細胞質の形態計測、およびAR+/AMACR-上皮核におけるARチャンネルの平均強度（例えば、免疫蛍光を示す組織画像のコンピュータ分析に基づいて生成）が含まれるフィーチャに基づく、前立腺摘除後の臨床的な失敗リスクを予測する予測モデルを提供する。モデルは、患者のデータセットを評価し、それにより、患者の臨床的な失敗リスクを評価することができる。

40

【0020】

別の実施形態では、1つまたは複数（例えば、全て、1つ、2つ、3つ、4つなど）の以下：生検グリーンスコア、リンパ節の関与、前立腺摘除グリーンスコア、組織画像由来の上皮細胞質の形態計測（例えば、上皮細胞質の平均強度）、組織画像由来の上皮核の形態計測（例えば、上皮核中のテクスチャの変動）、組織画像由来の間質の形態計測（例えば、間質内のテクスチャの変動）、および/またはラセマーゼ（AMACR）-陽性上皮細胞中のアンドロゲン受容体（AR）の強度（例えば、免疫蛍光を示す組織画像のコンピュータ分析に基づいて生成）が含まれるフィーチャに基づく、前立腺摘除後の臨床的

50

な失敗リスクを予測する予測モデルを提供する。モデルは、患者のデータセットを評価し、それにより、患者の臨床的な失敗リスクを評価することができる。

【0021】

別の実施形態では、複数の各時点での患者のデータを測定し、本発明の予測モデルによって評価するシステムおよび方法を提供する。患者の診断または治療は、各評価由来の結果の比較に基づき得る。かかる比較を、例えば、医師または他の個人による使用のためにコンピュータによるレポート出力にまとめることができる。例えば、医学的状態のインヒビター化合物についてのスクリーニングのためのシステムおよび方法を提供することができる。患者のための第1のデータセットを、臨床データ、分子データ、およびコンピュータ生成された形態計測データに基づく予測モデルによって評価することができる。試験化合物を患者に投与することができる。試験化合物の投与後、第2のデータセットを患者から得て、予測モデルによって評価することができる。第1のデータセットの評価の結果を、第2のデータセット由来の評価の結果と比較することができる。第1のデータセットに関する第2のデータセットの結果の変化により、試験化合物がインヒビター化合物であることを示すことができる。

10

【0022】

本発明のさらに別の態様では、医学的状態の出現を治療、診断、および/または予測するための試験キットを提供する。かかる試験キットを、病院、他の医療施設、任意の他の適切な場所に配置することができる。試験キットは、患者のデータ（例えば、臨床データ、分子データ、および/またはコンピュータ生成された形態計測データが含まれる）を受信し、患者のデータを予測モデル（例えば、試験キットのメモリ中にプログラミングする）と比較し、比較結果を出力することができる。いくつかの実施形態では、分子データおよび/またはコンピュータ生成された形態計測データを、試験キットによって少なくとも部分的に生成することができる。例えば、分子データを、患者の組織サンプルを受けた後に分析アプローチによって生成することができる。組織サンプルの電子画像の1つまたは複数のオブジェクトへのセグメント化、1つまたは複数のオブジェクトの1つまたは複数のオブジェクトクラス（例えば、間質、管腔、赤血球など）への分類、および1つまたは複数のオブジェクトクラスについての1つまたは複数の測定による形態計測データの決定を行うことによって形態計測データを生成することができる。いくつかの実施形態では、試験キットは、例えば、予測モデルへのアップデートを受信するための入力を含むことができる。いくつかの実施形態では、試験キットは、例えば、患者への請求および/または用途の追跡に有用なデータなどのように、データの別のデバイスまたは場所への転送のための出力を含むことができる。

20

30

【図面の簡単な説明】

【0023】

本発明の実施形態のより深い理解のために、添付の図面と共に以下の説明の後にリファレンスを示す。この図面中の引用符合は、明細書全体の符合に対応する。

【図1A】図1Aおよび1Bは、医学的状態の出現を治療、診断、または予測するために予測モデルを使用するシステムのブロック図である。図1Cは、予測モデルを生成するためのシステムのブロック図である。

40

【図1B】図1Aおよび1Bは、医学的状態の出現を治療、診断、または予測するために予測モデルを使用するシステムのブロック図である。図1Cは、予測モデルを生成するためのシステムのブロック図である。

【図1C】図1Aおよび1Bは、医学的状態の出現を治療、診断、または予測するために予測モデルを使用するシステムのブロック図である。図1Cは、予測モデルを生成するためのシステムのブロック図である。

【図2】図2は、予測モデルによって出力することができる患者の例示的結果を示す。

【図3】図3は、組織画像の処理に関する例示的ステージの流れ図である。

【図4】図4は、医学的状態のインヒビター化合物のスクリーニングに関する例示的ステージの流れ図である。

50

【図5】図5 aおよび5 bは、それぞれ画像のセグメント化および分類後の健康および異常な前立腺組織検体のグレースケールデジタル画像を示す。

【図6】図6は、前立腺癌再発を予測するためのモデルによって使用される種々の臨床的フィーチャ、分子的フィーチャ、およびコンピュータ生成された形態計測的フィーチャを示す。

【図7】図7 aおよび7 bは、2つの分子的フィーチャ、特に、アンドロゲン受容体（AR）およびCD34の存在を証明する染色組織検体を示す。

【図8】図8は、図6のフィーチャに基づいて予想される前立腺癌再発の経験について低リスク、中程度のリスク、または高リスクの患者の分類を示す Kaplan-Meier 曲線のグラフである。

【図9】図9は、前立腺癌再発を予測するためのモデルによって使用される種々の臨床的フィーチャ、分子的フィーチャ、およびコンピュータ生成された形態計測的フィーチャを示す。

【図10】図10は、図9のフィーチャに基づいて予想される前立腺癌再発の経験について低リスク、中程度のリスク、または高リスクの患者の分類を示す Kaplan-Meier 曲線のグラフである。

【図11】図11は、前立腺癌の全生存性を予測するためのモデルによって使用される種々の臨床的フィーチャ、分子的フィーチャ、およびコンピュータ生成された形態計測的フィーチャを示す。

【図12】図12は、図11のフィーチャに基づいて予想されるモデルによって予測される任意の原因に起因する死の低リスク、中程度のリスク、または高リスクの患者の分類を示す Kaplan-Meier 曲線のグラフである。

【図13】図13は、前立腺摘除した患者の後の浸潤性疾患を予測するためのモデルによって使用される種々の臨床的フィーチャおよびコンピュータ生成された形態計測的フィーチャを示す。

【図14】図14は、前立腺癌再発を予測するためのモデルによって使用される種々の臨床的フィーチャ、分子的フィーチャ、およびコンピュータ生成された形態計測的フィーチャを示す。

【図15】図15は、前立腺癌再発を予測するためのモデルによって使用される種々の臨床的フィーチャ、分子的フィーチャ、およびコンピュータ生成された形態計測的フィーチャを示す。

【図16】図16は、根治的前立腺摘除後の患者の臨床的な失敗を予測するためのモデルによって使用される種々の臨床的なフィーチャおよびコンピュータ生成された組織画像のフィーチャを示す。

【図17】図17は、前立腺癌再発を予測するためのモデルによって使用される種々の臨床的フィーチャおよびコンピュータ生成された組織画像のフィーチャを示す。

【図18】図18 a ~ e は、前立腺組織中の核 AR の分類のための複数の画像セグメント化を示す。

【図19】図19は、根治的前立腺摘除後の患者における臨床的な失敗を予測するための別のモデルによって使用される種々の臨床的フィーチャおよびコンピュータ生成された組織画像のフィーチャを示す。

【図20】図20は、カットポイント39を超えるかそれ未満のモデルスコア（39.5以上のスコアは総生存期間が短いと予測される）によって2群に分類されゲフィチニブで処置された非小細胞肺癌（NSCLC）を有する個体の生存を予測するためのモデルにおける患者51人についての生存曲線を示す線グラフである。

【発明を実施するための形態】

【0024】

（好ましい実施形態の詳細な説明）

本発明の実施形態は、医学的状態の出現の予想のための予測モデルにおいてコンピュータ生成された形態情報のみまたは臨床情報および/または分子情報との組み合わせを使用

10

20

30

40

50

する方法およびシステムに関する。例えば、本発明の1つの実施形態では、臨床情報、分子情報、およびコンピュータ生成された形態情報を使用して、前立腺癌の再発を予測する。他の実施形態では、本明細書中に提供された教示を使用して、他の医学的状態（例えば、他の疾患型（例えば、上皮新生物および混合新生物（乳房、結腸、肺、膀胱、肝臓、膵臓、腎臓の細胞、および軟組織が含まれる）など）の出現および1つまたは複数の療法（例えば、調合薬（*pharmaceutical drug*））に対する患者の応答性または非応答性を予測する。これらの予測を、医師または他の個人が使用して、例えば、患者に適切な一連の治療を選択し、および/または患者の医学的状態を診断することができる。

【0025】

本発明の1つの態様では、臨床的フィーチャ、分子のフィーチャ、およびコンピュータ生成された形態計測的フィーチャと医学的状態との間の相関を決定する分析ツール（サポートベクトルマシン（*SVM*）および/または神経回路網が含まれる）を提供することができる。相関するフィーチャは、容態の出現または再発を予測するために使用することができるモデルを形成することができる。例えば、分析ツールを使用して、医学的状態に関連するその転帰（例えば、癌再発までの期間）が少なくとも部分的に既知の患者のコホートについてのデータに基づいた予測モデルを生成することができる。次いで、モデルを使用して新規の患者のデータを評価し、新規の患者の医学的状態の出現を予測することができる。いくつかの実施形態では、3つのデータ型のうちのサブセットのみ（例えば、臨床データおよび形態計測データのみ）を分析ツールによって使用して、予測モデルを生成することができる。

【0026】

本発明の実施形態によって使用される臨床データ、分子データ、および/または形態計測データには、医学的状態の診断、治療、および/または予測に関連する任意の臨床データ、分子データ、および/または形態計測データが含まれ得る。予測モデルを生成するための前立腺癌の再発および生存との相関について分析したフィーチャは、例えば、表1、2、4および/または6に関連して以下に記載されている。少なくともいくつかのこれらのフィーチャ（例えば、上皮新生物および混合新生物）によって他の医学的状態（例えば、乳房、結腸、肺、膀胱、肝臓、膵臓、腎臓の細胞、および軟組織）の予測モデルの開発の基礎を得ることができると理解されるであろう。例えば、表1、2、4および/または6中の1つまたは複数のフィーチャを、いくつかの他の医学的状態を有する患者について評価し、次いで、分析ツールに入力して、フィーチャが医学的状態と相関するかどうかを決定することができる。モデルの医学的状態発生予測力を増加させるフィーチャを最終モデルに含めることができるのに対して、予測力を増加させない（または減少させる）フィーチャを、検討事項から除去することもできる。予測モデル開発の基礎としての表1、2、4および/または6中のフィーチャの使用により、医師、他の個人、および/または自動化処理装置（例えば、組織画像分析システム）のリソースを、転帰と相関する可能性がより高く、したがって、最終予測モデルで有用な患者のデータの取得に向けることができる。さらに、前立腺癌の再発および生存と相関すると判断されるフィーチャを、図6、9、および11に示す。これらのフィーチャを、前立腺癌の再発および/または生存を予測する最終モデルに直接含め、および/または他の医学的状態のための予測モデルの開発のために使用することができると理解されるであろう。

【0027】

形態計測データには、種々の構造および/またはスペクトル特性（例えば、組織検体）を示すコンピュータ生成されたデータが含まれ得る。1つの実施形態では、形態計測データには、間質、細胞質、上皮核、間質核、管腔、赤血球、組織人工物、組織バックグラウンド、またはその組み合わせの形態計測的フィーチャについてデータが含まれ得る。本発明の1つの実施形態では、組織画像から形態計測的フィーチャを測定するための組織画像分析システムを提供する。かかるシステムは、*Definiens Cellenger* ソフトウェアを使用する *MAGIC*（商標）システムであり得る。かかるシステムは、入

10

20

30

40

50

力としてH & E染色画像を受信し、画像中の病理学的オブジェクトについて種々の形態計測的フィーチャの測定値を出力することができる。画像から形態計測的フィーチャを得るためのシステムおよび方法に関するさらなる詳細は、図3に関して以下に記載されている。

【0028】

臨床的フィーチャは、年齢、人種、体重、身長、病歴、遺伝子型、および病状などの1つまたは複数の患者のデータを含むかこれらに基づき得る。ここで、病状は、臨床的および病理学的な病期のフィーチャおよび特に手作業で疾患過程について収集した任意の他の臨床的フィーチャをいう。一般に、臨床データは、医師によって患者および/または患者の組織または細胞の一連の試験中に収集される。臨床データには、特定の医学的状况により特異的であり得る臨床データも含まれ得る。例えば、前立腺癌の文脈では、臨床データには、前立腺特異抗原(P S A)の血中濃度、直腸の指診の結果、グリーソンスコア、および/または前立腺癌に対してより特異的であり得る他の臨床データを示すデータが含まれ得る。一般に、表1、2、4および/または6および/または図6、9および/または11中の任意のフィーチャ(すなわち、臨床的、形態計測的、および/または分子的なフィーチャ)を前立腺以外の医学的状况に適用する場合、前立腺に対してより特異的な表および/または図由来のフィーチャを考慮しないかもしれない。任意選択的に、問題の医学的状况に対してより特異的なフィーチャを、前立腺に特異的なフィーチャに代えることができる。例えば、他の組織学的に疾患特異的なフィーチャ/顕在化には、壊死領域(例えば、*in situ*での乳房の管癌)、サイズ、形状、および上皮細胞(例えば、乳房、肺)の領域パターン/分布、分化度(例えば、非小細胞肺癌を伴う扁平細胞分化(N S C L C、乳房および結腸の両方で認められる種々の腺癌に伴って認められるムチン産生)、細胞の形態計測的/微視的分布(例えば、乳癌の管裏層、N S C L C中の細気管支裏層)、ならびに炎症の程度および型(例えば、前立腺と比較して乳房およびN S C L Cが異なるフィーチャを有する)が含まれ得る。

10

20

【0029】

分子的フィーチャには、生体分子(核酸、ポリペプチド、サッカリド、ステロイド、および他の小分子または上記の組み合わせ(例えば、糖タンパク質およびタンパク質-R N A複合体)が含まれる)の存在、非存在、相対的な増加または減少、または相対的位置を示すデータを含むか、これらに基づき得る。これらの分子が測定される位置には、腺、腫瘍、間質、および/または他の位置が含まれ得、この位置は、特定の医学的状况に依存し得る。一般に、分子データを、一般的な分子生物学的技術および生化学的技術(サザンブロット、ウェスタンブロット、およびノーザンブロット、ポリメラーゼ連鎖反応(P C R)、免疫組織化学、および免疫蛍光)を使用して収集する。さらに、*in situ*ハイブリッド形成を使用して、分子生物学的フィーチャの相対存在量および位置の両方を示すことができる。組織の*in situ*ハイブリッド形成のための例示的方法およびシステムは、2003年7月21日出願され、発明の名称が「Methods and compositions for the preparation and use of fixed-treated cell-lines and tissue *in situ* fluorescence *in situ* hybridization」である上記援用の米国特許出願第10/624,233号(現在、米国特許第6,995,020号)に記載されている。

30

40

【0030】

図1Aおよび1Bは、患者の医学的状况の出現を予測するために予測モデルを使用する例示的システムを示す。例えば、*medical diagnostics lab*がリモートアクセスデバイスに関連した医師または他の個人の医学的決定を補助する場合、図1A中の配置を使用することができる。例えば、病院、他の医療施設、または他の適切な場所などの施設で用いる予測モデルを含む試験キットが提供される場合、図1B中の配置を使用することができる。

【0031】

50

図1Aに関して、予測モデル102を診断施設104に配置する。予測モデル102は、患者のデータを受信し、これを評価して患者の医学的状態の出現（例えば、再発）を予測し、評価の結果を出力するための任意の適切なハードウェア、ソフトウェア、またはその組み合わせを含むことができる。別の実施形態では、モデル102を使用して、特定の1つまたは複数の療法に対する患者の応答性を予測することができる。診断施設104は、インターネットサービスプロバイダー（ISP）108ならびに通信ネットワーク110および112を介してリモートアクセスデバイス106から患者のデータを受信し、評価のために予測モデル102にデータを入力することができる。遠隔地からの患者のデータの受信および評価のための他の配置も勿論可能である（例えば、電話線などの別の接続または通常郵便による）。遠隔地に存在する医師または個人は、任意の適切な様式で患者のデータを取得し、リモートアクセスデバイス106を使用してデータを診断施設104に転送することができる。いくつかの実施形態では、患者のデータを、診断施設104または他の施設によって少なくとも部分的に生成することができる。例えば、診断施設104は、リモートアクセスデバイス106または他のデバイスからデジタルバージョンのH&E染色画像を受信し、画像に基づいて患者の形態計測データを生成することができる。別の例では、実際の組織サンプルを診断施設104によって受信および処理して、形態計測データを生成することができる。他の例では、第三者が新規の患者の画像または組織を受信し、画像または組織に基づいて形態計測データを生成し、診断施設104に形態計測データを提供することができる。組織の画像および/またはサンプルから形態計測データを生成するための適切な画像処理ツールは、図3に関して以下に記載されている。

10

20

【0032】

診断施設104は、ISP108ならびに通信ネットワーク110および112または普通郵便または電話などの別の様式を介したリモートアクセスデバイス106への転送によってリモートアクセスデバイス106に関連する医師または個人に評価の結果を提供することができる。結果には、診断「スコア」（例えば、事象の再発までの予測時間などの医学的状態に関連する1つまたは複数の転帰を患者が経験する可能性の表示）、医学的状態と相関する予測モデル102によって分析された1つまたは複数のフィーチャを示す情報、予測モデルの感受性および/または特異性を示す情報、他の適切な診断情報、またはその組み合わせが含まれる。例えば、図2は、予測モデルによって出力することができる架空の患者についてのレポートの例を示す。図に示すように、レポートは、時間（月）（x軸）に対する患者の転帰の確率（例えば、前立腺癌の再発；すなわち、y軸）をマッピングしている。この例では、患者は、患者が高リスクカテゴリーに置かれた場合、スコア「520」を有する。医師または他の個人は、かかるレポートを使用して、より洗練された臨床診断腫瘍分類の決定を補助し、患者の再分類に有効な手段を開発し、最後に、各患者に対してより正確な（且つ適切な）治療オプションを生成することができる。レポートはまた、医師または個人が患者のリスクを患者に説明するのを補助することができるという点で有用であり得る。

30

【0033】

リモートアクセスデバイス106は、診断施設104（例えば、パーソナルコンピュータ、ワイヤレスデバイス（ラップトップコンピュータ、携帯電話、または携帯情報端末（PDA）など）、または任意の他の適切なリモートアクセスデバイスなど）からデータを転送および/または受信することができる任意のリモートデバイスであり得る。複数のリモートアクセスデバイス106を、図1Aのシステムに含めることができるが（例えば、対応する複数の遠隔地の複数の医師または他の個人がデータを診断施設104と通信するため）、描画が過剰に複雑になるのを回避するために、たった1つのリモートアクセスデバイス106を図1Aに含めた。診断施設104は、リモートアクセスデバイス106への、および/またはリモートアクセスデバイス106からの、伝達を受信および処理することができるサーバを含むことができる。かかるサーバは、コンピューティングハードウェアおよび/または記憶装置の異なる構成要素を含むことができるが、ソフトウェアアプリケーションまたはハードウェアとソフトウェアとの組み合わせであり得る。1つまたは

40

50

複数のコンピュータを使用してサーバを構成することができる。

【0034】

各通信リンク110および112は、任意の適切な有線または無線の通信経路（例えば、ローカルエリアネットワーク、広域ネットワーク、電話回線網、ケーブルテレビネットワーク、イントラネット、またはインターネットなど）であり得る。いくつかの適切な無線通信ネットワークは、移動通信（GSM）ネットワーク、時分割多元接続（TDMA）ネットワーク、符合分割多元接続（CDMA）ネットワーク、ブルートゥースネットワーク、または任意の他の適切な無線ネットワークのためのグローバルシステムであり得る。

【0035】

図1Bは、施設124で使用する本発明の予測モデルを含む試験キット122を提供するシステムを示す。施設124は、病院、医師のオフィス、または他の適切な場所であり得る。試験キット122は、患者のデータ（例えば、少なくとも1つの臨床データ、形態計測データ、および分子データ）を受信し、予測モデル（例えば、試験キットのメモリ中でプログラミングされた）を使用して患者のデータを評価し、評価の結果を出力するように構成された任意の適切なハードウェア、ソフトウェア、またはその組み合わせ（例えば、パーソナルコンピュータ）を含むことができる。例えば、試験キット122は、予測モデルの機能実施のためのコンピュータ実行可能命令を使用してコード化されたコンピュータ読み取り可能な媒体を含むことができる。予測モデルは、（例えば、図1C中のシステムなどの別のシステムまたはアプリケーションによって）事前に生成された所定のモデルであり得る。いくつかの実施形態では、試験キット122は、任意選択的に、例えば、組織サンプルまたは画像から形態計測的フィーチャに対応するデータを生成することができる画像処理ツールを含むことができる。適切な画像処理ツールは、図3に関して以下に記載されている。他の実施形態では、試験キット122は、例えば、入力デバイス（例えば、キーボード）または別のデバイスもしくは場所からの入力としてプレパッケージングされた形態計測的フィーチャのデータを受信することができる。試験キット122は、任意選択的に、例えば、予測モデルに対するアップデートを受信するための入力を含むことができる。試験キットはまた、任意選択的に、データ（患者への請求および/または用途の追跡に有用なデータなど）の主施設または他の適切なデバイスもしくは場所に転送するための出力を含むことができる。請求データは、例えば、試験キットによって評価された患者の医療保険情報（例えば、名称、保険会社、およびアカウント番号）を含むことができる。かかる情報は、例えば、試験キットの提供者が使用毎にキットに課金する場合、および/または提供者が保険会社に請求書を提出するために患者の保険情報が必要である場合に有用であり得る。

【0036】

図1Cは、例示的な予測モデル生成システムを示す。システムは、分析ツール132（例えば、サポートベクトルマシン（SVM）および/または神経回路網が含まれる）およびその転帰が少なくとも部分的に知られた患者のデータベース134を含む。分析ツール132は、データベース134由来のデータと医学的状态との間の相関を決定するための任意の適切なハードウェア、ソフトウェア、またはその組み合わせを含むことができる。図1C中のシステムはまた、例えば、デジタル化バージョンのH&E染色組織画像、実際の組織サンプル、またはその両方に基づいて形態計測データを生成することができる画像処理ツール136を含むことができる。ツール136は、例えば、そのデータがデータベース134に含まれる既知の患者の形態計測データを生成することができる。適切な画像処理ツール136は、図3に関して以下に記載されている。

【0037】

データベース134は、臨床的フィーチャ、形態計測的フィーチャ、分子的フィーチャ、またはその組み合わせのデータなどの任意の適切な患者データを含むことができる。データベース134はまた、患者が疾患再発を経験したかどうかおよび経験した時期などの患者の転帰を示すデータを含むことができる。例えば、データベース134は、医学的状态の再発を経験した患者のデータなどの患者の打ち切りなしのデータ（uncensored

10

20

30

40

50

e d d a t a) (すなわち、その転帰が完全に知られている患者のデータ)を含むことができる。データベース134は、選択的またはさらに、1回または複数回の医師への追跡調査来院において疾患再発の兆候が認められなかった患者のデータなどの患者の打ち切りデータ(すなわち、その転帰が完全に知られていない患者のデータ)を含むことができる。分析ツール132による打ち切りデータの使用は、予測モデルの生成に利用可能なデータ量を増大させることができ、したがって、モデルの信頼性および予測力を有利に改善することができる。打ち切りデータおよび打ち切りなしのデータの両方を使用することができるサポートベクトルマシン(SVM)および神経回路網(NNCi)の例を、以下に記載する。

【0038】

1つの実施形態では、分析ツール132は、サポートベクトルマシン(SVM)を含むことができる。かかる実施形態では、ツール132は、好ましくは、打ち切りデータ上のサポートベクトル回帰(SVRc)を実施することができるSVMを含む。同時係属米国特許出願第10/991,240号に記載のように、SVMが打ち切りデータを使用することができるSVM内で使用するための新規の修正損失関数/ペナルティ関数をSVRc中に提供する。データベース134由来の既知の患者の臨床的フィーチャ、分子的フィーチャ、および/または形態計測的フィーチャを含むデータをSVMに入力して、予測モデルのパラメータを決定することができる。パラメータは、入力フィーチャの相対的重要性を示すことができ、SVMが既知の患者の転帰を予測する能力を最大にするように調整することができる。フィーチャの医学的状態との相関を決定するためのSVMの使用に関するさらなる詳細を、(5)および(6)に記載する。

【0039】

分析ツール132によるSVRcの使用は、患者の状態を示す情報の多次元非線形ベクトルをデータベース134から取得することを含むことができる。ここで、少なくとも1つのこのベクトルは対応する患者に関する事象の出現時間の表示を欠く。次いで、分析ツール132は、ベクトルを使用して回帰を行ってカーネルベースのモデルを生成し、それにより、情報のベクトル中に含まれる少なくともいくつかの情報に基づいて事象までの時間の予測に関連する出力値を得る。分析ツール132は、打ち切りなしのデータを含むベクトルについてツール132によって使用される損失関数と異なる打ち切りデータを含む各ベクトルの損失関数を使用することができる。打ち切りデータサンプルを差分的に操作することができる。これは、打ち切りデータサンプルが「片方の情報」のみしか提供できないからである。例えば、生存期間の予測では、打ち切りデータサンプルは、典型的には、所与の時間内に事象が起こらなかったことのみを示し、たとえ起こったとしても、所与の時間後に起こる場合を示さない。

【0040】

打ち切りデータのために分析ツール132によって使用された損失関数は、以下であり得る。

【0041】

【数1】

$$Loss(f(\mathbf{x}), y, s=1) = \begin{cases} C_s^*(e - \epsilon_s^*) & e > \epsilon_s^* \\ 0 & -\epsilon_s \leq e \leq \epsilon_s^* \\ C_s(\epsilon_s - e) & e < -\epsilon_s \end{cases}$$

ここで、 $e = f(\mathbf{x}) - y$; および $f(\mathbf{x}) = W^T(\mathbf{x}) + b$ は、フィーチャ空間Fにおける線形回帰関数である。ここで、Wは、Fのベクトルであり、 (\mathbf{x}) はFのベクトルに対する入力xをマッピングする。

【0042】

対照的に、打ち切りなしのデータのためにツール132によって使用される損失関数は、以下であり得る。

10

20

30

40

50

【 0 0 4 3 】

【 数 2 】

$$\text{Loss}(f(\mathbf{x}), y, s=0) = \begin{cases} C_n^*(e - \varepsilon_n^*) & e > \varepsilon_n^* \\ 0 & -\varepsilon_n \leq e \leq \varepsilon_n^* \\ C_n(\varepsilon_n - e) & e < -\varepsilon_n \end{cases}$$

ここで、 $e = f(\mathbf{x}) - y$

および

$$C_n^* \text{ および } C_n \text{ は } C_n^* \text{ および } C_n \text{ である。}$$

10

【 0 0 4 4 】

上記説明では、 W および b は、最適化問題を解決することによって得られる。その一般式は以下である。

【 0 0 4 5 】

【 数 3 】

$$\min_{W, b} \frac{1}{2} W^T W$$

$$\text{s.t. } y_i - (W^T \phi(x_i) + b) \leq \varepsilon$$

20

$$(W^T \phi(x_i) + b) - y_i \leq \varepsilon$$

しかし、この方程式は、凸最適化問題が常に実現可能であると仮定しており、これは、真ではないかもしれない。さらに、回帰推定中の小さな誤差を許容することが望ましい。これらが損失関数を $SVRC$ のために使用する理由である。損失により、回帰推定をおこなう余地がいくらかできる。理想的には、構築されたモデルは全ての結果を正確に計算するが、これは実行不可能である。損失関数は、理想からの一定範囲の誤差が可能であり、この範囲はスラック変数 ξ および ξ^* ならびにペナルティ C によって制御される。理想から逸脱しているが、 ξ および ξ^* によって定義される範囲内である誤差を計数するが、その寄与を C によって軽減する。インスタンスの誤差が多いほど、ペナルティが高くなる。インスタンスの誤差がより少ないほど（理想に近い）、ペナルティが少ない。誤差と共にペナルティが増加するこの概念によって勾配が生じ、 C はこの勾配を制御する。種々の損失関数を ε - insensitive 損失関数のために使用することができる一方で、一般方程式を、以下のように変形する。

30

【 0 0 4 6 】

【 数 4 】

$$\min_{W, b} P = \frac{1}{2} W^T W + C \sum_{i=1}^l (\xi_i + \xi_i^*)$$

$$\text{s.t. } y_i - (W^T \Phi(x_i) + b) \leq \varepsilon + \xi_i$$

$$(W^T \Phi(x_i) + b) - y_i \leq \varepsilon + \xi_i^*$$

$$\xi_i, \xi_i^* \geq 0, \quad i=1 \dots l$$

40

本発明の ε - insensitive 損失関数について（打ち切りデータおよび打ち切りなしのデータに適用した異なる損失関数を使用）、この方程式は、以下のようになる。

【 0 0 4 7 】

【数 5】

$$\begin{aligned} \min_{\mathbf{w}, b} P_c &= \frac{1}{2} \mathbf{W}^T \mathbf{W} + \sum_{i=1}^l (C_i \xi_i + C_i^* \xi_i^*) \\ \text{s.t.} \quad y_i - (\mathbf{W}^T \Phi(\mathbf{x}_i) + b) &\leq \varepsilon_i + \xi_i \\ (\mathbf{W}^T \Phi(\mathbf{x}_i) + b) - y_i &\leq \varepsilon_i^* + \xi_i^* \\ \xi_i^{(*)} &\geq 0, \quad i=1 \dots l \end{aligned}$$

ここで、 $C_i^{(*)} = s_i C_s^{(*)} + (1-s_i) C_n^{(*)}$
 $\varepsilon_i^{(*)} = s_i \varepsilon_s^{(*)} + (1-s_i) \varepsilon_n^{(*)}$

10

最適化基準は、その y 値が を超えることによって $f(x)$ と異なるデータポイントにペナルティを科す。スラック変数 および $*$ は、それぞれ正の偏差および負の偏差についての過剰な偏差サイズに対応する。このペナルティ機構は、2つの成分（一方は打ち切りなしのデータ（すなわち、右打ち切りでない）および他方は打ち切りデータを有する。両成分を、ここに、 - 無視損失関数と呼ばれる損失関数形態で示す。

【0048】

打ち切りデータにおけるサポートベクトル回帰 (SVRC) を実施するためのシステムおよび方法に関するさらなる詳細は、2004年11月17日出願された上記で援用された米国特許出願第10/991,240号および2003年11月18日出願された米国特許仮出願第60/520,939号に記載されている。

20

【0049】

別の実施形態では、分析ツール132は、神経回路網を含むことができる。かかる実施形態では、ツール132は、好ましくは、打ち切りデータを使用することができる神経回路網を含む。さらに、神経回路網は、好ましくは、関連モデル (NNci) を訓練するために、一致指数 (CI) の近似 (例えば、微分) に実質的にしたがって目的関数を使用する。生存率分析の成績指標としてCIが長く使用されているが (7)、神経回路網を訓練するためのCIの使用は以前に提案されていなかった。これまで、目的関数の訓練としてCIを使用することは、CIが微分不可能であり、勾配ベースの方法によって最適化できないという点で困難である。2005年2月25日出願された同時係属米国特許出願第11/067,066号 (現在、米国特許第7,321,881号) に記載のように、この障害を、目的関数としてCIの近似値を使用することによって克服することができる。

30

【0050】

例えば、分析ツール132が前立腺癌再発を予測するために使用される神経回路網を含む場合、神経回路網は、前立腺癌再発に関するその転帰が出力を得るために少なくとも部分的に既知である患者のコホートについての入力データを処理することができる。神経回路網への入力のために選択された特定のフィーチャを、上記SVRC (例えば、分析ツール132のサポートベクトルマシンを使用して実行される) の使用または別の適切なフィーチャ選択過程の使用によって選択することができる。ツール132の誤差モジュールは、出力と入力データに対応する所望の出力との間の誤差 (例えば、予測される転帰と患者の既知の転帰との間の差異) を決定することができる。次いで、分析ツール132は、神経回路網の成績を評価するためにCIの近似に実質的にしたがって目的関数を使用することができる。分析ツール132は、目的関数の結果に基づいて神経回路網の重み付きコネクション (weighted connection) (例えば、フィーチャの相対的重要性) に適応することができる。フィーチャと予測された転帰との相関を調整するための神経回路網の重み付きコネクションの適応に関するさらなる詳細は、(8) および (9) に記載されている。

40

【0051】

一致指数を、以下の形態：

【0052】

【数 6 - 1】

$$CI = \frac{\sum_{(i,j) \in \Omega} I(\hat{t}_i, \hat{t}_j)}{|\Omega|}$$

ここで、

$$I(\hat{t}_i, \hat{t}_j) = \begin{cases} 1: \hat{t}_i > \hat{t}_j \\ 0: \text{上記でなければ} \end{cases},$$

10

で示すことができ、これは、それぞれ、患者 i および j についての予後評価

【0053】

【数 6 - 2】

\hat{t}_i

と

【0054】

【数 6 - 3】

\hat{t}_j

20

との間の一対比較に基づき得る。この例では、 \hat{t}_i は、以下の条件を満たす全患者対 { i , j } からなる。

【0055】

・両患者 i および j は再発を経験し、患者 i の再発時間 t_i は、患者 j の再発時間 t_j よりも短いか、

・患者 i のみが再発を経験し、 t_i は患者 j の追跡調査来院時間 t_j よりも短い。

CI の分子は、神経回路網によってより早く再発すると予測される患者が実際により早く再発する時間数を示す。分母は、所定の条件を満たす患者対の総数である。

【0056】

一般に、CI が増加する場合（好ましくは、最大になる場合）、モデルはより正確である。したがって、好ましくは CI を実質的に最大にするか CI の近似により、モデルの成績は改善される。本発明の 1 つの実施形態は、以下に示す CI の近似を提供する。

30

【0057】

【数 7 - 1】

$$C = \frac{\sum_{(i,j) \in \Omega} R(\hat{t}_i, \hat{t}_j)}{|\Omega|}$$

ここで、

$$R(\hat{t}_i, \hat{t}_j) = \begin{cases} (-(\hat{t}_i - \hat{t}_j - \gamma))^n : \hat{t}_i - \hat{t}_j < \gamma \\ 0 : \text{上記でなければ} \end{cases},$$

40

(式中、 $0 < \gamma < 1$ および $n > 1$)。

【0058】

【数 7 - 2】

$R(\hat{t}_i, \hat{t}_j)$

を、

【0059】

50

【数 7 - 3】

$$I(-\hat{t}_i, -\hat{t}_j)$$

の近似と見なすことができる。

【0060】

結果を改善するために経験的に示した本発明によって提供されるCIの別の近似を以下に示す。

【0061】

【数 8 - 1】

$$C_{\omega} = \frac{\sum_{(i,j) \in \Omega} -(\hat{t}_i - \hat{t}_j) \cdot R(\hat{t}_i, \hat{t}_j)}{D},$$

10

ここで、

【0062】

【数 8 - 2】

$$D = \sum_{(i,j) \in \Omega} -(\hat{t}_i - \hat{t}_j)$$

は、規格化因子である。ここで、各

【0063】

【数 8 - 3】

$$R(\hat{t}_i, \hat{t}_j)$$

を、

【0064】

【数 8 - 4】

$$\hat{t}_i$$

と

【0065】

【数 8 - 5】

$$\hat{t}_j$$

との間の差異によって重みを付ける。C（またはC）の最小化は、中の各サンプル対を、

【0066】

【数 8 - 6】

$$\hat{t}_i - \hat{t}_j > \gamma$$

を満たし、それによって

【0067】

【数 8 - 7】

$$I(\hat{t}_i, \hat{t}_j) = 1$$

が得られるように模索する。

【0068】

中の対の出力間の差異が縁より大きい場合、このサンプル対は目的関数への寄与を停止するであろう。この機構は、モデルの訓練中のデータの過剰適合を有効に克服し、好ましくは、中により多くのサンプル対を移動させて

【0069】

50

【数 8 - 8】

$$\hat{i}_i - \hat{i}_j > \gamma$$

を満たすことのみ注目する。訓練サンプルの影響を、訓練中の一対比較にしたがって適応的に調整する。R 中の切除縁陽性 が汎化性能の改良に好ましいことに留意のこと。言い換えれば、神経回路網のパラメータを、全患者データを入力した後に CI を計算することによって訓練中に調整する。次いで、神経回路網は、目的関数を最小化するためにパラメータを調整し、それにより、CI が最大化する。上記で使用するように、過剰適合は、一般に、神経回路網の複雑さをいう。具体的には、ネットワークが非常に複雑な場合、ネットワークは「ノイズの多い」データに反応するであろう。過剰適合は、データの訓練範囲をはるかに超えて容易に予測することができるという点で危険である。

10

【0070】

神経回路網を訓練するための CI の近似に実質的にしたがって目的関数を使用するためのシステムおよび方法に関するさらなる詳細は、2005年2月25日に出願された米国特許出願第 11 / 067, 066 号（現在、米国特許第 7, 321, 881 号）、2004年2月27日に提出された米国特許仮出願第 60 / 548, 322 号および 2004年6月4日に提出された同第 60 / 577, 051 号に記載されている。

【0071】

図 3 は、適切な画像処理ツールの例示的機能の流れ図である。図 3 中の機能は、主に、画像中の病理学的オブジェクトの分類（例えば、細胞質、管腔、核、間質、バックグラウンド、人工物、および赤血球のようなオブジェクトの分類）のための組織画像のセグメント化に関する。1つの例では、画像処理ツールは、SPOT Insight QE カラーデジタルカメラ（KAI 2000）を使用して 20 倍の組織画像を取り込み、1600 × 1200 ピクセルの画像を生成する光学顕微鏡を含むことができる。画像を、Tiff 形式で 24 ビット / ピクセルの画像として保存することができる。かかる装置は例示のみを目的とし、本発明の範囲を逸脱することなく任意の他の適切な画像取り込み装置を使用することができる。画像処理ツールはまた、取り込まれた画像をセグメント化および分類し、次いで、オブジェクトの形態計測的フィーチャを測定するための任意の適切なハードウェア、ソフトウェア、またはその組み合わせを含むことができる。1つの実施形態では、画像処理ツールは、例えば、上記の種々の病理学的オブジェクトのセグメント化および分類を行い、これらのオブジェクトの種々の形態計測的フィーチャを測定するのに適した市販の Definiens Cellenger Developer Studio (v. 4.0) を含むことができる。Definiens Cellenger 製品に関するさらなる詳細は、(10) に記載されている。画像処理ツールは、オブジェクトの種々の形態計測的フィーチャ（スペクトルベースのフィーチャ（赤色、緑色、青色（RGB）チャンネル特性（平均値、標準偏差など）、位置、サイズ、周囲長、形状（非対称性、楕円適合度（elliptic fit）など）が含まれる）、隣接するオブジェクトに対する空間的關係および強度の關係（明暗差）を測定することができる。画像処理ツールは、画像中のあらゆる同定された病理学的オブジェクトのあらゆるインスタンスについてこれらのフィーチャを測定し、例えば、予測モデル 102（図 1A）、試験キット 122（図 1B）、または分析ツール 132（図 1C）による評価のためにこれらのフィーチャを出力することができる。任意選択的に、画像処理ツールはまた、測定した各フィーチャの画像についての全統計的概要を出力することができる。分類された病理学的オブジェクトの形態計測的フィーチャの測定に関するさらなる詳細は、表 1 および 2 に関して以下に記載されている。以下は、図 3 に示す画像処理ツールの機能の説明である。

20

30

40

【0072】

最初のセグメント化。第 1 ステージでは、画像処理ツールは、画像（例えば、H & E 染色組織マイクロアレイ（TMA）画像または全組織切片の H & E）をオブジェクトとして公知の小さな連続的ピクセル群にセグメント化することができる。これらのオブジェクトを、色類似性および形状規則性に基づいて連続領域を見出す領域拡張法によって得ること

50

ができる。オブジェクトのサイズを、いくつかのパラメータの調整によって変化させることができる(11)。このシステムでは、ピクセルよりもむしろオブジェクトが、典型的には、最も小さな処理単位である。したがって、全ての形態計測的フィーチャの計算および操作を、オブジェクトに関して行うことができる。例えば、画像に閾値を適用する場合、オブジェクトのフィーチャ量(feature values)を閾値に供する。結果として、オブジェクト内の全てのピクセルを同一クラスに割り当てる。1つの実施形態では、オブジェクトのサイズを、最高詳細レベルで10~20ピクセルであるように制御することができる。このレベルに基づいて、より低いレベルでのより小さなオブジェクトからのより大きなオブジェクトの形成によってその後のより高く且つより粗いレベルを構築する。

10

【0073】

バックグラウンド抽出。最初のセグメント化後、画像処理ツールは、強度閾値および凸包を使用してバックグラウンド(スライドの透明領域)から画像組織コアをセグメント化することができる。強度閾値は、画像ピクセルを2つのクラス「組織コア」および「バックグラウンド」に分類する強度値である。閾値以上の強度値を有する任意のピクセルを、「組織コア」ピクセルとして分類し、そうでなければ、ピクセルを「バックグラウンド」ピクセルと分類する。幾何学的オブジェクトの凸包は、このオブジェクトを含む最も小さな凸集合(多角形)である。セットSは、凸面であり、いつでも2つのポイントPおよびQがSの内側に存在する場合、ラインセグメントPQもS中に存在する。

粗セグメント化。次のステージでは、画像処理ツールは、フォアグラウンド(例えば、TMAコア)を核および余白に対応するラフ領域(rough region)に再セグメント化することができる。例えば、H&E染色画像中の核のフィーチャの主な特性は、残りの病理学的オブジェクトと比較して青色に染色されることである。したがって、赤色チャンネルと青色チャンネル(R-B)の強度値の差異を、識別されるフィーチャとして使用することができる。特に、最初のセグメント化工程で得たあらゆる画像オブジェクトについて、平均の赤色ピクセル強度値と青色ピクセル強度値との間の差異を決定することができる。長さ/幅比を使用して、オブジェクトが核領域として分類すべきかどうかを決定することもできる。例えば、(R-B)フィーチャ閾値および長さ/幅閾値を下回るオブジェクトを、核領域と分類することができる。同様に、緑色チャンネル閾値を使用して、余白として組織コア中にオブジェクトを分類することができる。組織間質は、赤色で占められる。強度差異d、「赤色比」 $r = R / (R + G + B)$ 、および画像オブジェクトの赤色チャンネル標準偏差Rを使用して、間質オブジェクトを分類することができる。

20

30

【0074】

余白の分類。粗セグメント化ステージでは、余白領域は、画像中の管腔(病理学的オブジェクト)および人工物(破壊組織領域)の両方に対応し得る。より小さな余白オブジェクト(100ピクセル未満の領域)は、通常、人工物である。したがって、画像処理ツールは、これらを人工物として分類するための領域フィルタ(area filter)を適用する。

【0075】

核の脱融合および分類。粗セグメント化ステージでは、核領域は、しばしば、いくつかの実際の核を含む連続的な融合領域として得られる。さらに、核領域は、周辺の誤分類された細胞質も含み得る。したがって、これらの融合核領域は、各核を得るために脱融合される必要があり得る。

40

【0076】

画像処理ツールは、核を脱融合するために2つの異なるアプローチを使用することができる。第1のアプローチは、形状拘束(円形)下で核領域を構成する画像オブジェクトを融合する領域拡張法に基づき得る。このアプローチは、融合が深刻でない場合にうまく行くように決定されている。

【0077】

深刻な融合の場合、画像処理ツールは、管理学習(supervised learn

50

ing) に基づいた異なるアプローチを使用することができる。このアプローチは、専門家(病理学者)による核領域の手作業のラベリングを含む。ラベリングした核に属する画像オブジェクトのフィーチャを使用して、統計分類子(statistical classifier)をデザインすることができる。

【0078】

1つの実施形態では、フィーチャ空間の次元数を減少させるために、以下の2つの分類子を使用して、訓練組に対してフィーチャ選択を行うことができる: Bayesian分類子およびk最近傍分類子(12)。1つ取って置き法(leave-one-out method)(13)をクロスバリデーションのために使用することができ、逐次前方探索法(sequential forward search method)を使用して、最高のフィーチャを選択することができる。最後に、2つのBayesian分類子を、それぞれ、1および5に等しいフィーチャ数を使用してデザインすることができる。クラス条件付き分布は、対角共分散行列を有するガウス分布であると推測することができる。

10

【0079】

いくつかの実施形態では、入力画像は、以下の異なる種の核を含むことができる: 上皮核、線維芽細胞、基底核、内皮核、アポトーシス核、および赤血球。上皮核数が典型的には腫瘍範囲の分類における重要なフィーチャと見なされるので、上皮核数は上皮核を他と区別するのに重要であり得る。画像処理ツールは、検出された核の以下の2つのクラスへの分類によってこれを行うことができる: 形状(偏心性(eccentricity))およびサイズ(領域)のフィーチャに基づいた上皮核および「残り」。

20

【0080】

本発明の画像のセグメント化および分類に関するさらなる詳細は、2004年11月17日に出願された上記で援用された米国特許出願第10/991,897号および2003年11月17日に出願された米国特許仮出願第60/520,815号および2004年3月12日に出願された同第60/552,497号に記載されている。

【0081】

上記のように、画像処理ツールは、ツールによる画像中のオブジェクトのセグメント化および分類後に種々の形態計測的フィーチャを測定することができる。これらの形態計測的フィーチャは、1つまたは複数の特性および/または統計を示すことができる。オブジェクトの特性には、スペクトル特性(例えば、カラーチャンネルの平均値、標準偏差、および輝度)および構造/形状特性(例えば、領域、長さ、幅、稠密度、密度)の両方が含まれ得る。統計値には、最大値、最小値、平均値、および標準偏差が含まれ得、画像オブジェクトの各特性を計算することができる。表1および表2(本明細書中に添付)は、本発明にしたがって測定することができる形態計測的フィーチャの種々の例を示す。これらの表中の形態計測的フィーチャを、これらのフィーチャによって測定された種々の特性および/または統計値を示す従来法を使用して命名する。表1および表2に示す特定の命名規則を、上記の市販のDefinensソフトウェア製品から取り入れる。したがって、当業者に理解されるであろう。

30

【0082】

表1および表2中に示したコンピュータ生成された形態計測的フィーチャは例示のみを目的とし、任意のコンピュータ生成された形態計測的フィーチャを本発明の範囲を逸脱することなく使用することができるという理解されるであろう。例えば、表1および表2は、異なる形態計測的フィーチャセットを含む。表2中のフィーチャセットの減少および修正(すなわち、表1のフィーチャと比較した減少および修正)は、表1を含む研究を行った時間からの前立腺癌の再発および生存の分野におけるさらなる実験に起因した。特に、さらなる実験により、転帰と相関する可能性がより高いかもしれないフィーチャ型に関するさらなる洞察が得られた。発明者は、実験の継続および/または他の適切なハードウェア、ソフトウェア、またはその組み合わせの使用によってこれらおよび他の医学的状態と相関し得る種々の他のコンピュータ生成されたフィーチャセット(例えば、表1中のフィーチャ

40

50

ャのサブセット(表10および表11を参照のこと)または表2中のフィーチャのサブセット)が得られると予想する。

【0083】

表1および表2に関して、フィーチャ「Lumen.StdDevAreaPx1」、「管腔」は、画像オブジェクトの型を示し、「StdDev」は同定された管腔の全てのインスタンスを使用して計算されるべき統計値(標準偏差)を示し、「AreaPx1」は統計値によって評価されるべきオブジェクトインスタンス(ピクセル数としての領域)のフィーチャを示す。画像処理ツールは、画像中で事前にセグメント化および分類された全オブジェクトについての形態計測的フィーチャを測定することができる。例えば、画像処理ツールは、オブジェクトの形態計測的フィーチャ(「バックグラウンド」、「細胞質」、「上皮核」、「管腔」、「間質」、「間質核」、および「赤血球」が含まれる)を測定することができる。「バックグラウンド」は、組織で占められていないデジタル画像の部分を含む。「細胞質」は細胞の細胞質をいい、無定形領域(例えば、例えば、H&E染色組織の画像中の上皮核を囲む桃色領域)であり得る。「上皮核」は、腺単位(glandular unit)の上皮細胞/管腔細胞および基底細胞内に存在する核をいい、細胞質に囲まれた「円形」オブジェクトとして出現する。「管腔」は、上皮細胞によって分泌物が堆積する中心の腺状空間をいい、上皮細胞に囲まれた閉ざされた白色領域として出現する。時折、管腔は、前立腺液(典型的には、H&E染色組織中に桃色で出現する)または他の「破片」(例えば、マクロファージ、死細胞など)で満たされ得る。管腔、上皮細胞質、および核を合わせて、腺単位を形成する。「間質」は、前立腺組織の構造を維持する異なる密度を有する結合組織の一形態をいう。間質組織は、腺単位の間が存在し、H&E染色組織中で赤色から桃色として出現する。「間質核」は、細胞質(線維芽細胞)を含まないか最少量で含む細長い細胞である。このカテゴリーには、内皮細胞および炎症細胞も含まれ得、癌が存在する場合、間質内に上皮核を見出すこともできる。「赤血球」は、通常、血管(動脈または静脈)内に存在する小さな赤色の円形の物体であるが、組織全体の分散状態でも見出され得る。

10

20

【0084】

以下の表中の「C2EN」は、細胞質に対する核領域の比である。上皮細胞がより未分化/悪性になれば、より多くの領域を核が占め、比率がより大きくなる。「EN2SN」は、デジタル組織画像中に存在する間質細胞に対する上皮細胞の比率または相対量である。「L2Core」は、組織内に存在する管腔の数または領域である。グリーンソン分類がより高ければ、癌がより高い侵襲性を示し、したがって、管腔存在量がより少なくなる。一般に、これは、癌が生じる場合に無制御の方法で上皮細胞が複製され、管腔が上皮細胞で満たされるようになるからである。

30

【0085】

本発明の1つの態様では、医学的状态(例えば、疾患)のインヒビター化合物をスクリーニングするためのシステムおよび方法を提供する。図4は、本発明の1つの実施形態のインヒビター化合物のスクリーニングに関する例示的ステージの流れ図である。ステージ402では、患者の第1のデータセット(1つまたは複数の臨床データ、形態計測データ、および分子データが含まれる)を得ることができる。試験化合物を、ステージ404で患者に投与することができる。ステージ404後、ステージ406で第2のデータセットを患者から得ることができる。第2のデータセットは、第1のデータセット中に含まれる同一のデータ型(すなわち、フィーチャ)を含んでも、含まなくてもよい。ステージ408では、第2のデータセットを、第1のデータセットと比較することができ、ここで試験化合物の投与後の第2のデータセットの変化は、試験化合物がインヒビター化合物であることを示す。データセット比較のステージ408は、第1のデータセットの入力に応答する本発明の予測モデルによって生成された出力を第2のデータセットの入力に応答する予測モデルによって生成された出力と比較することを含むことができる。例えば、インヒビター化合物は所与の薬物であってよく、本発明は、薬物が医学的状态の医学的処置として有効であるかどうかを決定することができる。

40

50

【 0 0 8 6 】

医学的状態の予測に対する本発明の実施形態の種々の例示的適用をここに記載するであろう。第1の例では、本発明の1つの実施形態は、前立腺癌の再発を予測するために臨床データおよび形態計測データを使用した。第2の例では、本発明の1つの実施形態は、前立腺癌の再発および全生存性を予測するために臨床データ、形態計測データ、および分子データを使用した。第3の例では、本発明の1つの実施形態を使用して、患者の前立腺摘除後の浸潤性疾患の出現を予測することができる。第4の例では、本発明の1つの実施形態を使用して、肝臓毒物学を予測した。第5、第6、および第8の例では、本発明の実施形態を使用して、前立腺癌の再発を予測した。第7および第9の例では、本発明の実施形態を使用して、前立腺摘除後の臨床的な失敗を予測した。第10の例では、本発明の1つの実施形態を使用して、ゲフィチニブで処置した非小細胞肺癌（NSCLC）の個体の生存率を予測した。

10

【 0 0 8 7 】

前立腺癌の概要

前立腺癌は、米国男性の主な死亡原因であり、2004年に新規に診断された症例は230,000と予想され、約30,000人が死亡している。PSAを使用した血清ベースのスクリーニングの使用の拡大により、医師は、前立腺に局在している初期段階（すなわち、T1a-c、T2）または領域に拡大しているが検出される比率が小さい転移期に前立腺癌を検出することができる。初期検出および初期診断の報告された利点は、一連の処置の選択において患者および泌尿器科医の両方に非常に大きな重荷を課した。腫瘍の大部分が無痛であり、最少の介入（すなわち、「静観」）が必要である一方で、他の腫瘍はより浸潤性が高く、早期介入（すなわち、放射線療法/ホルモン療法/全身補助療法/臨床試験の配置）が推奨されるので、最初の治療介入を選択する場合、正確な診断は重要である。さらに、静観と根治的前立腺摘除とを比較する無作為化試験では、手術がわずかに有利であり（前立腺摘除後の死亡率が6.6%減少）、個人に合わせた患者のケアをガイドするためにより良好な患者の層別化が必要であることが示唆される（14）。

20

【 0 0 8 8 】

PCaの自然史は、診断時に患者が直面する困難を再度強調する（15）。早期前立腺癌が局所療法を使用して治癒可能である場合でさえ、約25~40%の男性がPSA/生化学的再発（BCR）を起こすであろう。さらにより厄介なことには、前立腺癌が再発した男性は、PSA/BCRの約8年後（BCR後平均8年；中央値5年）に依然として転移し得る。これにより、この患者群の治療レジメン（そのBCRまでの時間および転移傾向の予測の両方）における早期のこの患者群の同定がその全生存に最重要であることが示唆される。不運なことに、既存の予測モデルはその正確性が制限され、その腫瘍病理に関して特定の患者のために個別化されていない。種々の遺伝的变化、環境的变化、および生活様式の変化がPCaの病理発生に関与しているにもかかわらず、現在、所与の患者の転帰を予測することができる単一の生物学的経路、遺伝子変異、または臨床的バイオマーカーは存在しない。根治的前立腺摘除が再度一般的になってから21年後およびPSA使用が広まってから15年後、泌尿器科医は、局所性疾患の治療によって最良に臨床疾患がなくなるか、全生存する患者を依然として区別できない。

30

40

【 0 0 8 9 】

臨床的フィーチャデータにのみ基づいた予後ノモグラムは、実際、臨床状態および転帰を有用に予測するが、精度および普遍性の両方を改良する必要がある（16）。本発明の実施形態は、PSA/BCR前立腺摘除後の予後モデルの正確性を首尾よく改良するための「システム病理学（systems pathology）」アプローチを提供する。これは、高度に正確かつ統合的な予測モデルを作製するための患者自身の腫瘍サンプルの「個別化した」表示（細胞および微小解剖学的形態計測特性、臨床プロフィール、および分子マーカーが含まれる）を示す。専門分野の使用により、PSA再発の予想のための高度に正確な予測モデルが開発されている。これらの試みにより、予測モデルおよび予後モデルの生成におけるシステム病理学の有用性が立証された。さらに、分析により、制限さ

50

れた臨床的変数、分子バイオマーカー、および組織形態計測的フィーチャの組を誘導することができ、これを、指定された臨床結果に基づいて最適な患者の処置プランを構築するために泌尿器科医/病理学者によって使用される予測試験に含めることができることが示されている。PSAの再発に関連した選択された分子的フィーチャにより、成長因子シグナル伝達(下記のアンドロゲン受容体(以後、「AR」)による)および細胞結合脈管形成(CD34による)の機構についての収束性の役割が示唆される。CD34は、上皮細胞上に存在し、身体の脈管を裏打ちしている膜貫通糖タンパク質である。これらの所見および前立腺癌の進行の予測に及ぼす潜在的影響をより深く理解するためのさらなる研究は進行中である。モデルの開発および精度において重要ないくつかの新規の組織記述子(descriptor)に加えて、高度に正確で非主観的な定量的グリーンスコアを一部示す選択された画像のセグメント化および形態計測的特性にも留意した。グリーンスコアリングシステムに関する定義された形態計測的フィーチャには、腺構造の全外観の一部、上皮細胞の形状およびサイズ(細胞質成分)、上皮細胞核、および間質中で混合された単一上皮細胞の証明が一部含まれる。

10

20

30

40

50

【0090】

アンドロゲン受容体タンパク質(AR)は、これらのホルモンが精巣のライディッヒ細胞によって合成された後に天然に存在する男性ホルモン(テストステロンおよびその5還元代謝産物であるジヒドロテストステロン)を受け取る。特に、合成後、これらのホルモンは体内を循環し、ARに結合する。受容体ARを介して作用するアンドロゲンは、胎児中の男性生殖器および副生殖腺の発達、思春期男性における男性化および成長、ならびに成人における男性生殖力および生殖機能の維持を刺激する。アンドロゲン受容体は、他のステロイドホルモン受容体と共に、特異的遺伝子配列との相互作用を介して遺伝子転写を調節するトランス作動性転写調節タンパク質ファミリーを構成する。

【0091】

前立腺癌に関するARの研究により、癌細胞中のアンドロゲン受容体の存在と成長のための男性ホルモン刺激の依存との間に正相関が存在し得ることが示唆されている。例えば、Sovak et al.の米国特許第6,472,415号は、早期における前立腺癌の成長はアンドロゲン駆動であり、少なくとも一過性にアンドロゲン枯渇によって停止することができることを提案している。French et al.の米国特許第6,821,767号は、医師による前立腺癌の診断的評価におけるアンドロゲン受容体アッセイの使用が可能なARの種々の測定方法を提案している。しかし、これらの研究は、本明細書中に開示の前立腺癌の出現を予測する自動化モデルと併せたAR測定の使用を提案していなかった。

【実施例】

【0092】

(実施例1)前立腺癌再発の予測

臨床データおよび形態計測データ

最初に500ほどの多数の生形態計測的フィーチャを、Definiens Cellengerソフトウェアに基づくMAGIC組織画像分析システムを使用して、各前立腺組織から抽出した。潜在的に有用なフィーチャが無視されるのを回避するために、生フィーチャの全セットを不可知的に選択した。しかし、これらの全形態計測的フィーチャは、平等に有益でありそうになく、全フィーチャセットに基づいて構築された予測モデルは、「次元の呪い(curse of dimensionality)」によって予測成績が不十分である可能性が高いであろう(13)。それにより、次元削減手順を適用し、8つの形態計測的フィーチャセットを最終的に選択した。

【0093】

根治的前立腺摘除を受けた前立腺癌患者コホートから患者153人のサブセットに基づいて研究を行った。術後の測定可能な前立腺特異抗原(PSA)を使用して、前立腺癌の再発(生化学的再発(BCR)ともいう)を定義した。術後に患者を追跡した。最後の来院におけるその再発状態およびその追跡期間を記録し、右打ち切りデータセットを生成し

た。グリーソンスコアを、術前に生検検体から測定し、術後の切除した前立腺を使用して測定した。以下の4つの特異的な臨床的測定（すなわち、フィーチャ）を本研究で考慮した：（1）生検グリーソン分類、（2）生検グリーソンスコア、（3）術後グリーソン分類、および（4）術後グリーソンスコア。

【0094】

臨床的に導かれたグリーソンスコアフィーチャ由来の形態計測的フィーチャを個別に分析して、PSA/BCR再発の確率およびPSA/BCR再発までの期間を予測した。次いで、画像およびグリーソンスコア（フィーチャ）を組み合わせ、再発を確立し、再発までの期間を予測した。この共同のフィーチャセットによって達成された予測精度の改善は、画像フィーチャが実際にさらなる情報を提供し、それにより、再発予測率および全予測モデルを強化したことを示した。

10

【0095】

この患者コホートが右打ち切り転帰データを有していたので、生存分析モデルは、再発予測のために構築されなければならない。異なるデータ型における潜在的バイアスを回避するために、以下の2つの生存分析法を使用した：1）コックス回帰モデル（17）および2）上に記載されており、サポートベクトルマシンに適用されるSVRc。5倍クロスバリデーションを使用して評価した一致指数を使用して、モデルの予測精度を測定した（13）（18）。

【0096】

両方法を、以下の3つのデータセットに適用した：（1）グリーソンスコア臨床的フィーチャのみ；（2）選択された形態計測的フィーチャのみ；および（3）形態計測的フィーチャとグリーソンスコア臨床的フィーチャとの組み合わせ。実験結果を、表3に列挙する。

20

【0097】

本実施例で選択された臨床的フィーチャは、BXGGTOT、BXGG1、GGTOT、およびGG1であり、形態計測的フィーチャを、以下に関して選択した：上皮核（Epithelial.Nuclei.MaxCompactness）、バックグラウンド（Background.StdDevAreaPx1）、および管腔（Lumen.MaxBorderLengthPx1、Lumen.MinRadiusofsmallestenclosure、Lumen.StdDevBorderLengthPx1、Lumen.SumBorderlengthPx1、Lumen.StdDevAreaPx1、およびLumen.MinCompactness）。より具体的には、本実施例では、管腔の領域、境界長、および形状（稠密度）に関連する形態計測的フィーチャを、疾患の進行に相関させるように決定した。管腔がより小さく、且つより稠密度が高いほど、癌が進行する可能性が高い。実際、より浸潤性の高い癌（グリーソン分類4および5）では、管腔が組織からほとんどまたは完全に消失すると予想され得る。上皮核領域に対する上皮核の長さおよび幅の比としてDefinens Cellengerソフトウェアによって稠密度を計算した場合、上皮核の稠密度の形態計測的フィーチャは癌の進行と相関すると判断された。これは、間質への上皮核の浸潤が癌進行につれて増加する（すなわち、進行癌を有する組織は、典型的には、多数の上皮核を含む）からであり得る。本実施例の転帰と相関すると判断されたバックグラウンドベースの形態計測的フィーチャにより、本分析で使用した実際の組織コアサイズを測定した。

30

40

【0098】

【表3】

表3-予測精度の比較

	グリーソン	画像	グリーソン +画像
Cox	0.6952	0.6373	0.7261
SVRc	0.6907	0.7269	0.7871

50

表3によれば、形態計測的フィーチャの予測性能はグリーンスコアに匹敵し、形態計測的フィーチャとグリーンスコアとの組み合わせによってより高い予測率が達成される。この高い予測率により、組織画像分析システムによって抽出された形態計測的フィーチャが実際にグリーンスコアを超えるさらなる情報を提供することが確認される。したがって、形態計測の使用により、全再発予測を強化することができる。

【0099】

(実施例2)

前立腺癌再発および全生存の予測

臨床データ、形態計測データ、および分子データ

前立腺特異抗原 (PSA) の再発をそれぞれ 88% および 87% の予測精度で首尾よく予測する 2 つの研究を行った。臨床的フィーチャ、分子的フィーチャ、および形態計測的フィーチャと機械学習との組み合わせにより、ロバストプラットフォームを作製した。このプラットフォームは、患者の診断、治療管理、および予後診断で広く適用されている。注目の転帰が任意の原因による死亡であった前立腺癌患者の全生存を予測するために第 3 の研究を行った。

10

【0100】

根治的前立腺摘除を受けた 539 患者のコホートの研究を、前立腺摘除検体から構築した高密度組織マイクロアレイ (TMA) を組み込みつつ行った。ヘマトキシリンおよびエオシン (H&E) 染色組織切片を使用して形態計測研究を行い、分子生物学的決定因子を、免疫組織化学 (IHC) を使用して評価した。PSA の再発および全生存の両方のための予測モデルは、管理された多変量学習によって選択されたフィーチャセットから導いた。各ドメイン中の完全な非欠測データ (complete non-missing data) を有する患者を、打ち切りデータを操作するために開発した回帰のためのサポートベクトルマシン (SVRc) を使用して評価した。モデルの予測性能を、リスク群を定義するために使用した生成スコアを伴う一致指数 (CI) を使用して評価した。

20

【0101】

132 患者のコホートを使用して、PSA 再発を精度 88% で予測した 41 個のフィーチャ (17 個の臨床的フィーチャ、14 個の分子的フィーチャ、および 10 個の形態計測的フィーチャを含む) を選択した。268 患者のコホートでは、10 個のフィーチャ (3 臨床的フィーチャ、1 個の分子的フィーチャ、および 6 個の形態計測的フィーチャ) により、PSA 再発を精度 87% で予想可能なことを見出され、さらに、14 個のフィーチャ (2 個の臨床的フィーチャ、1 個の分子的フィーチャ、および 11 個の形態計測的フィーチャ) により、全生存を精度 80% で予測可能なことを見出された。ログランク検定を使用して、リスク群の間で腫瘍再発および死亡の有意差が認められた ($p < 0.0001$)。

30

【0102】

本研究により、機械学習で分析される臨床的変数、分子マーカー、および組織学を組み合わせた新規のシステムアプローチの使用によって前立腺癌の再発傾向の予測が前進的に改善されることが明らかとなる。

【0103】

患者の臨床的フィーチャ。根治的前立腺摘除を受けた 539 患者のコホートを研究した。17 個の臨床的フィーチャ (以下の表 4 に示す) を、患者の年齢、術前 PSA、およびグリーンソン分類が含まれる非特定の (de-identified) 患者情報を使用して遡及的に収集した。

40

【0104】

【表 4】

表4. 収集した臨床的フィーチャ

フィーチャ	説明
age	年齢(歳)
race	人種
prepsa	前立腺特異抗原(ng/dl)
tnm	TNM臨床病期
uicc	UICC臨床病期
dre	直腸の指診で触診可能
ln	リンパ節の状態
svi	精囊の浸潤
margins	+/- 切除縁
ece	被膜外側の腫瘍
bxgg1	優位生検グリーソン分類
bxggtot	生検グリーソンスコア
gg1	優位術後グリーソン分類
ggtot	術後グリーソンスコア
prsltd	二倍体、四倍体、異数性
pp_sphas	S期の倍数性における細胞の比率
pp_frac	倍数性増殖画分

10

組織マイクロアレイ (TMA) を、前立腺摘除検体の選択されたブロックから構築した。各検体由来の直径 0.6 mm の組織コアを、各レシピエントパラフィンブロックについて三連で無作為に整列させた (Beecher Instruments, Silver Spring, MD)。切片 (5 μm) のこれらの TMA ブロックを、形態計測分析および免疫組織化学 (IHC) 分析のために使用される荷電ポリリジンコーティングスライド上に置いた (以下を参照のこと)。

20

【0105】

臨床的フィーチャの欠測値を、全フィーチャを含む柔軟なさらなる回帰モデルを使用して帰属させて転帰と無関係にフィーチャの欠測値を評価し、完全な臨床データ (帰属後)、形態計測データ、および分子データならびに非欠測転帰の情報を有する患者のみをさらに研究した。研究 1 (概念実証) に有効なサンプルサイズは、患者 132 人からなる。目的の一次分類は、患者が前立腺癌手術後に再発するかどうかについてであった。PSA > 0.2 ng/mL で 2 つの連続した評価を行った患者を、前立腺癌が再発したと見なした。患者がその最後の来院以後に再発しなかったか患者の転帰が最近の来院以後に未知である場合 (すなわち、追跡失敗)、患者の転帰を打ち切りデータと見なした。再発までの期間を、根治的前立腺摘除から PSA (生化学的) 再発までの期間 (月) と定義した。

30

【0106】

研究 1 由来の患者 132 人のうちの 129 人を含む最初の患者 539 人のコホート由来の患者 268 人を使用して研究 2 を行った。TMA コア由来の H & E 画像を使用する代わりに、根治的前立腺摘除由来の全切片を分析した。研究 3 は、同一の患者 268 人のコホートを試験したが、全生存を予測するために使用し、注目の転帰は任意の原因による死亡であった。

【0107】

画像分析および形態計測研究。各患者の組織コアまたは全切片から回収された元の腫瘍組織の代表的領域をデジタル化し、H & E 染色スライドを使用して分析した。20 倍の光学顕微鏡によって SPOT Insight QE カラーデジタルカメラ (KAI 2000) を使用して画像を取り込んだ。80% 超が腫瘍を含む領域のみを、最適な画像セグメント化および定量分析のために選択した。

40

【0108】

分子分析。12 のバイオマーカー (サイトケラチン 18 (管腔細胞)、サイトケラチン 14 (基底細胞)、CD45 (リンパ球)、CD34 (内皮細胞)、CD68 (マクロファージ)、Ki67 (増殖)、PSA (hK-3、カリクレイン)、PSMA (成長受容体)、サイクリン D1 (細胞周期)、p27 (細胞周期)、アンドロゲン受容体 (内分泌)、および Her-2/neu (シグナル伝達) が含まれる) のパネルを、標準的な発色

50

免疫組織化学を使用して全部で7つのTMAブロックに適用した。全抗原についてプレッシャークッカー中で0.01Mクエン酸緩衝液(pH6)を30分間使用して抗原回復を行った。かかる過程に関する例示的方法およびシステムは、上記で援用され、発明の名称が「Methods and compositions for the preparation and use of fixed-treated cell-lines and tissue in fluorescence in situ hybridization」である2003年7月21日出願された米国特許出願第10/624,233号(現在、米国特許第6,995,020号)に記載されている。一次抗体(表5に示す)を、0.1%Tweenを含むTris緩衝化生理食塩水中で希釈し、4で16時間アプライし、その後1000倍希釈の二次抗体(Vector)でピオチン化した。

10

【0109】

【表5】

表5. 抗体のリスト

バイオマーカー	クローン
Ki-67	クローンki-67(DAKO)
サイトケラチン18	クローンDC-10(Novocastra)
CD45	クローンX16/99
CD68	クローン514H2(Novocastra UK)
CD34	クローンQBEnd 101(DAKO)
AR	クローンAR27(Novocastra)
サイトケラチン14	クローンLL002(Novocastra)
サイクリンD1	クローンP2D11F11
PSA	クローンPA05(Neomarkers)
PSMA	クローンZMD.80(Zymed) ^P
p27	クローンDCS72(Oncogene)
Her-2/neu	KIT DAKO ^P

20

^Pポリクローナル、残りはモノクローナル

ネガティブコントロールスライドに、一次抗体として正常マウス血清(DAKO)を投与した。スライドをHarrisヘマトキシリンで対比染色し、2人の独立した病理学者によって観察し、全ての矛盾を第3の病理学者によって解決した。539人の全患者およびその各三連のコア由来の記録したIHCデータは、調査中の特定の抗原について染色した細胞の比率および強度(0~3+)を含んでいた。適用可能な場合、これらの2つの測定を組み合わせ、特定のバイオマーカーについての染色指数を作製した(以下の表6は分子的フィーチャの例示的リストを示す)。AR(アンドロゲン受容体)、CK14(サイトケラチン14)、サイクリンD1、PSA(前立腺特異抗原)、PSMA(前立腺特異的膜抗原)、p27、およびHer2/neuの染色指数を計算する一方で、残りのマーカー(すなわち、Ki67、CK18(サイトケラチン18)、CD45、CD68)を、所与の強度を有する陽性細胞の比率に基づいて評価した。これらのバイオマーカーを、以下にさらに記載する。染色指数は0~300の範囲であり、以下のように計算した: $1 * (\text{バイオマーカーについて強度} 1 + \text{の細胞染色陽性の比率}) + 2 * (\text{バイオマーカーについて強度} 2 + \text{の細胞染色陽性の比率}) + 3 * (\text{バイオマーカーについて強度} 3 + \text{の細胞染色陽性の比率})$ (式中、細胞染色陽性の比率は、計算した100個の細胞あたり同定された陽性細胞数をいう)。この染色指数に関するさらなる詳細は、(19)に記載されている。かかる染色指数は例示のみを目的とし、本発明の範囲を逸脱することなく、分子的フィーチャの任意の他の測定方法を使用することができる。

30

40

【0110】

上記のバイオマーカーの考察では、p27は、サイクリン-CDK複合体に結合して細胞周期をG1期で停止させるサイクリン依存性キナーゼインヒビターと呼ばれる細胞周期レギュレーターファミリーに属する。バイオマーカーp27は、アポトーシスを促進し、いくつかの組織の最終分化で役割を果たすと改訂されている。免疫組織化学により、核p27発現の喪失は、より悪性の表現型に関連する。Her2/neuは受容体チロシンキナーゼのEGFRファミリーのメンバーであり、一定のヒト癌の病理発生において重要な

50

役割を果たす。細胞膜における免疫組織化学による Her 2 / neu の過剰発現は、より悪性の乳癌型に関連している。Ki 6 7 は、種々の強度で核を染色し、問題の腫瘍サンプルの増殖指数 (p r o l i f e r a t i v e i n d e x) を評価するか細胞活性を測定するために使用される多数の増殖マーカーの 1 つである。CD 4 5 は、リンパ球 (T 細胞、B 細胞、NK 細胞など) などの免疫細胞になる運命にある細胞を同定するために使用される細胞表面抗原である。強度は、その分布 / 存在および他の組織要素との関連ほど重要でないと考えられる。CD 6 8 は、リソソームと密接に関連する細胞質抗原である。これは単球分化カスケードを介して発現されるが、通常、単球よりマクロファージで強度が高い。

【 0 1 1 1 】

【 表 6 】

10

表6. 分子的フィーチャ

フィーチャ	説明
atki67t1	強度領域 1 中の Ki-67 (腫瘍)
atki67t2	強度領域 2 中の Ki-67 (腫瘍)
atki67t3	強度領域 3 中の Ki-67 (腫瘍)
atki67p1	強度領域 1 中の Ki-67 (PIN)
atki67p2	強度領域 2 中の Ki-67 (PIN)
atki67p3	強度領域 3 中の Ki-67 (PIN)
atki67a1	強度領域 1 中の Ki-67 (腺)
atki67a2	強度領域 2 中の Ki-67 (腺)
atki67a3	強度領域 3 中の Ki-67 (腺)
atc18t3	サイトケラチン 18 (腫瘍)
atcd45t3	CD45 (腫瘍)
atcd68t3	CD68 (腫瘍)
atcd34p	CD34 (PIN)
atcd34s	CD34 (間質)
atcd34t	CD34 (腫瘍)
atcd34tp	CD34 (腫瘍/PIN)
atcd34ts	CD34 (腫瘍/間質)
atcd34ps	CD34 (PIN/間質)
atc18p3	サイトケラチン 18 (PIN)
atcd45p3	CD45 (PIN)
atc18a3	サイトケラチン 18 (腺)
atcd45a3	CD45 (腺)
arsi	AR (腫瘍) 染色指数
c14si	サイトケラチン 14 (腫瘍) 染色指数
cd1si	サイクリン D1 (腫瘍) 染色指数
psasi	PSA (腫瘍) 染色指数
psmasi	PSMA (腫瘍) 染色指数
p27si	p27 (腫瘍) 染色指数
her2si	Her-2/neu (腫瘍) 染色指数
arpsi	AR (PIN) 染色指数
c14psi	サイトケラチン 14 (PIN) 染色指数
cd1psi	サイクリン D1 (PIN) 染色指数
psapsi	PSA (PIN) 染色指数
psmapsi	PSMA (PIN) 染色指数
p27psi	p27 (PIN) 染色指数
her2psi	Her-2/neu (PIN) 染色指数
arasi	AR (腺) 染色指数
c14asi	サイトケラチン 14 (腺) 染色指数
cd1asi	サイクリン D1 (腺) 染色指数
psaasi	PSA (腺) 染色指数
psmaasi	PSMA (腺) 染色指数
p27asi	p27 (腺) 染色指数
her2asi	Her-2/neu (腺) 染色指数

20

30

40

分析的および統計的研究。以下の 3 つの研究を行った：患者 1 3 2 人を使用した最初の概念実証分析 (研究 1) および患者 2 6 8 人を使用した長期調査 (研究 2 および研究 3) 。研究 1 および研究 2 の両方では、分析は、以下の 2 つの工程からなる：将来の根治的

50

立腺摘除患者における生化学的 (P S A) 再発を予測するためのモデル使用を最終目的とする、 P S A 再発を予測するフィーチャの同定およびこれらのフィーチャに基づいたモデルの構築。研究 3 の目的は、フィーチャを同定し、前立腺摘除後の全生存を予測するためのモデルを構築することである。上記型の打ち切りデータについてのサポートベクトル回帰 (S V R c) を使用して、これらの各研究で得られたモデルを構築した。

【 0 1 1 2 】

モデルの予測精度を、一致指数 (C I) を使用して評価した。打ち切りデータの転帰の取り扱いにおいて、これは、最適の測定法である。一致指数は、以下の基準のいずれか 1 つを満たす 2 人の無作為に選択した患者の予後スコアの間の一対比較に基づく：両患者において第 1 の患者の事象および事象時間が第 2 の患者より短いことを経験したか、第 1 の患者のみにおいて事象および第 1 の患者の事象時間が第 2 の患者の追跡時間よりも短いことを経験した。 C I は、モデル由来のより高い予後スコアを有する患者がより低いスコアを有する患者より短い時間内に事象を経験し、 R O C 曲線下の領域 (A U C) と強く関連する確率を評価する。他の測定基準を使用して、予測モデルの能力を測定することもできる。例えば、診断の評価で感度および特異性を使用することができる。別の例として、機会のみが、例えば、層間の認められた差異を担う確率を示す「 p 値」を使用することができる (例えば、図 8、10、および 12 を参照のこと)。したがって、 p 値がより低いと転帰に対して真の統計的関連がある可能性がより高くなる。典型的には、標準は、 0 . 0 5 以下の任意の p 値が統計的に有意であることである。

10

【 0 1 1 3 】

研究 1。本分析では、上記 S V R c モデルを、臨床データ、分子データ、および形態計測データに連側的に適用した。ここで、臨床的フィーチャは、分子データにつながる S V R c を介した「グリーディ・フォワード (g r e e d y - f o r w a r d) 」フィーチャ選択 (「 F S 」) 法のためのアンカーとしての役割を果たす。この工程後、アンカーとして臨床データと選択された分子的フィーチャとの組み合わせを使用して、形態計測データに対する第 2 の S V R c グリーディ・フォワードフィーチャ選択法を行った。最後の工程は、最終モデルを誘導するための臨床的フィーチャ、選択された分子的フィーチャ、および選択された形態計測的フィーチャとの組み合わせにおけるグリーディ・バックワード選択法の実施を含んでいた。フィーチャ選択中、フィーチャをモデル中に入力 (または保持) するかどうかを決定するための基準は、そのフィーチャの存在 (または非存在) によって一致指数が増加する (予測情報が増加する) かどうかに基づいた。

20

30

【 0 1 1 4 】

内部検証および外部検証の両方を使用して予測精度についてモデルを評価した。5 倍クロスバリデーションを使用して内部検証を行った。外部検証を行うために、患者コホートから一連の患者試験組を作製し、予測された転帰を、一致指数を介してこれらの患者の実際の転帰と比較した。この 2 レベルの検証デザインの適用では、患者のサブセットを患者レコードのフルセットから無作為に選択し、残りの患者のみを使用し、前述の手順を使用して予測モデルを構築した。次いで、非公開のレコードを使用して訓練されたモデルに適用し、予想精度を得た。これら 2 つの工程を B 回繰り返し、予測率 B を得た。ここで、最終予測率は平均であった。最終モデルのために選択されたフィーチャは、作製された異なるモデル B で十分な時間のようなフィーチャであった。

40

【 0 1 1 5 】

選択されたフィーチャセットを使用して、一致指数の直接的最大化によって神経回路網モデルを構築した。特に、上記型の神経回路網 (N N c i) を使用した。そのネットワークを、実質的に一致指数の近似にしたがって目的関数を使用して訓練した。最終モデルの出力を使用して、 P S A 再発についての各患者の将来のリスクを評価した。

【 0 1 1 6 】

研究 2。本研究の目的は、研究 1 と同一であった。しかし、異なるフィーチャの選択および検証手順を使用した。アンカーアプローチの代わりに、全フィーチャを P S A 再発 (一致指数によって測定) までの時間とのその関連によって格付けし、一定の所定の閾値 (

50

CI > 0.60) をパスしたフィーチャを選択した。フィーチャの画像化数を本発明者らの専門家によって減少した後にこれを行い、次いで、これらのフィーチャを一連の n - フィーチャモデル (例えば、1 - フィーチャ、2 - フィーチャ、3 - フィーチャなど) で評価した。フィーチャ前進選択手順を使用して、各 n - フィーチャモデルの一致指数を最大化するフィーチャを、次の n + 1 - フィーチャモデルで使用した。CI が所定の閾値によって改善することができないと、この過程を終了した。次いで、バックワードフィーチャ選択過程を使用して、CI を増加させる目的でフィーチャを除去した。任意のフィーチャの除去によって CI が改善されなかった場合、この過程を終了した。

【0117】

簡潔なブーストラップ技術を、フィーチャ選択のために使用した。このアプローチでは、患者を復元抽出して訓練組として使用した一方で、モデルを選択していない患者に関して評価した。比較として、このフィーチャ選択法を、Kattan et al に付与された米国特許第 6,409,664 号 (その全体が本明細書中で参考として援用される) に記載の Kattan 術後ノモグラムで見出されたフィーチャのみを使用して行った。最終モデルの出力を使用して、それぞれの未来の患者の PSA 再発リスクを評価した。

10

【0118】

研究 3。本研究の目的は、研究 2 で分析した同一のコホートおよびフィーチャならびに同一のフィーチャ選択法を使用して全生存を予測するフィーチャを同定することであった。最終モデルの出力を使用して、任意の原因によるそれぞれの未来の患者の死亡リスクを評価した。

20

【0119】

結果

一般的なアプローチは、システム病理学 (形態計測的分析と、分子署名と、患者の臨床プロフィールとの組み合わせ) を適用して、前立腺摘除後の前立腺癌患者の状態のコホートにおいて PSA 再発および全生存のための予測モデルを構築することであった。研究 1 由来の臨床病理学的フィーチャのみを標準的な Cox モデル分析で使用した場合、PSA 再発の予測精度は 59% でしかなかったことに留意することは重要である。形態計測的フィーチャおよび分子的フィーチャの SVRC との統合後に予測精度のレベルが 88% に増加した。以下の項は、どのようにしてこの改良が達成されたかを記載する。

30

【0120】

研究 1。このコホート中の患者 132 人について、診断における年齢の中央値は 63 歳 (最年少: 40 歳、最年長 81 歳) であり、術前 PSA の中央値は 8.2 ng/dl (最小値: 1.1、最大値: 81.9) であった。前立腺摘除サンプルに基づいて、32% がグリーンスコア 7 未満であり、60% がグリーソン 7 であり、残りの 8% が 7 超であった。患者 69 人 (52%) が pT2N0M0 であり、患者 40 人 (30%) が pT3aN0M0 であり、残りの患者 23 人 (18%) が pT3bN0M0 または pT1-3N+ であった (表 7 は、3 つの研究についての臨床的特徴のまとめリストを含む)。

【0121】

【表 7 - 1】

表 7. 臨床情報

40

	研究 1	研究 2 および 3
N	132	268
年齢 (歳)		
平均	62	62
中央値	63	63
範囲	40-81	40-81

【0122】

【表 7 - 2】

	研究1	研究2および3	
人種			
コーカサス	120 (90.9%)	241 (89.9%)	
ヒスパニック	8 (6.1%)	12 (4.5%)	
アフリカ系アメリカ人	2 (1.5%)	9 (3.4%)	
未知	2 (1.5%)	6 (2.2%)	
術前PSA (ng/dl)			
平均	12.2	10.8	
中央値	8.2	7.8	
範囲	1.1 - 81.9	0.9 - 81.9	
TNM期			
pT2N0	69 (52.3%)	157 (58.6%)	10
pT3aN0	40 (30.3%)	72 (26.9%)	
pT3bN0	13 (9.8%)	22 (8.2%)	
pT1-3N+	10 (7.6%)	17 (6.3%)	
UICC期			
T1a < 5%	0 (0.0%)	1 (0.3%)	
T1b ≥ 5%	0 (0.0%)	1 (0.3%)	
T1c (明確でも視覚可能でもない)	49 (37.1%)	112 (41.8%)	
T2a ≤ ½葉	23 (17.4%)	58 (21.7%)	
T2b ≤ 1葉	27 (20.5%)	45 (16.8%)	
T2c 両葉	23 (17.4%)	34 (12.7%)	
T3a 片側ECE	8 (6.1%)	15 (5.6%)	
T3c SV+	2 (1.5%)	2 (0.8%)	
DREの結果			
不明確	56 (42.4%)	118 (44.0%)	20
明確	76 (57.6%)	150 (56.0%)	
リンパ節の関与			
陰性	121 (91.7%)	250 (93.3%)	
陽性	11 (8.3%)	18 (6.7%)	
精囊の関与			
なし	113 (85.6%)	236 (88.0%)	
あり	19 (14.4%)	32 (12.0%)	
切除縁			
陰性	108 (81.8%)	217 (81.0%)	
陽性	24 (18.2%)	51 (19.0%)	
囊外の関与			
なし	70 (53.0%)	159 (59.3%)	
あり	62 (47.0%)	109 (40.7%)	
優位生検グリーソン分類			
1	0 (0.0%)	1 (0.4%)	30
2	24 (18.2%)	43 (16.0%)	
3	85 (64.4%)	184 (68.7%)	
4	22 (16.7%)	38 (14.2%)	
5	1 (0.7%)	2 (0.8%)	
生検グリーソンスコア			
2	0 (0.0%)	1 (0.4%)	
3	0 (0.0%)	0 (0.0%)	
4	6 (4.6%)	7 (2.6%)	
5	27 (20.5%)	56 (20.9%)	
6	41 (31.1%)	97 (36.2%)	
7	48 (36.4%)	90 (33.6%)	
8	7 (5.3%)	13 (4.9%)	
9	3 (2.3%)	4 (1.5%)	40
優位術後グリーソン分類			

【 0 1 2 3 】

【表 7 - 3】

	研究1	研究2および3
2	3 (2.3%)	20 (7.5%)
3	98 (74.2%)	201 (75.0%)
4	31 (23.5%)	47 (17.5%)
術後グリーソンスコア		
5	6 (4.6%)	21 (7.8%)
6	36 (27.3%)	86 (32.1%)
7	79 (59.9%)	148 (55.2%)
8	10 (7.6%)	12 (4.5%)
9	1 (0.8%)	4 (0.4%)
倍数性		
二倍体	74 (56.1%)	145 (54.1%)
四倍体	54 (40.9%)	115 (42.9%)
異数性	4 (3.0%)	8 (3.0%)
S期の倍数性率(%)		
平均	2.3	2.4
中央値	1.1	1.1
範囲	0.0 - 63.8	0.0 - 66.4
倍数性画分の比率		
平均	3.4	3.5
中央値	2.6	2.4
範囲	0.0 - 20.0	0.0 - 20.0

10

25人(15%)の患者がPSA再発を経験した一方で、残りの患者(85%)は打ちきられた。打ち切られた患者について、追跡期間の中央値は60.8ヶ月(すなわち、5年を少し超過する)であった。PSA再発までの全時間の中央値に届かなかった。PSA再発が予測される全17の臨床的フィーチャを選択し、そのうちで最も有益なフィーチャを以下に注釈する(モデルによって選択された臨床病理学的フィーチャおよび時間数): 生検グリーソン分類(112)、人種(112)、UICC臨床病期(110)、倍数性(110)、およびDREの結果(109)。

20

【0124】

画像分析および形態計測研究。図5aおよび図5bは、本発明にしたがってセグメント化および分類した後に得たそれぞれ健康および異常な前立腺組織のデジタル化画像を示す。例示のために、組織中の種々の病理学的オブジェクトをラベリングした。画像分析ソフトウェアによって全部で496個の形態計測的フィーチャ(表1に示す)を得た。

30

【0125】

496個の形態計測的フィーチャのうち、図6に示すPSA再発を予測する10個の形態計測的フィーチャを選択した。以下の病理学的オブジェクト(赤血球、上皮核、管腔、間質、細胞質、および組織バックグラウンド)に関連して以下の形態計測的フィーチャ(フィーチャの隣の括弧内の数字は、最終モデルの生成中に転帰と相関するフィーチャが何回選択されたかを示す)を選択した: 赤血球の最小長(ピクセル)(20)、上皮核の最大稠密度(17)、最小包囲(enclosure)の管腔最小半径(14)、上皮核最小幅(ピクセル)(11)、間質最大密度(10)、管腔最小境界長(ピクセル)(10)、上皮核最小標準偏差チャンネル2(10)、最小包囲の上皮核最大半径(10)、境界長の細胞質標準偏差(ピクセル)(10)、および領域のバックグラウンド標準偏差(ピクセル)(10)。より詳細には、本実施例では、赤血球の長さ、管腔の最小包囲の半径および境界長、細胞質の境界長、間質の密度(例えば、間質によって覆われた領域の平方根をその半径で除した値)、およびバックグラウンド領域の形態計測的フィーチャを、転帰と相関するように決定した。稠密度、幅、グリーンチャンネル値、および上皮核の最小包囲の半径の形態計測的フィーチャ(例えば、オブジェクトと同一領域を有する楕円を作製し、次いで、この楕円が上皮核を完全に囲むまで拡大し、元の楕円の半径に対する最も小さな包囲楕円の比を計算する)も、転帰と相関するように決定した。

40

【0126】

これらの相関のうちの少なくともいくつかについての種々の可能な理由は、実施例1と併せて上に記載されている。例えば、上皮核の稠密度の形態計測的フィーチャは、円周バ

50

ターンの上皮細胞の「背中合わせの」性質を反映し得る。それにより、腺および管腔の形成/分化の喪失が示唆され、したがって、より高いグリーソン分類（すなわち、より高い疾患の進行）と一致するであろう。また、管腔の最小包囲の半径の形態計測的フィーチャは、管腔の全サイズに関連し、グリーソン分類が増加するにつれて劇的に減少および縮小する。

【0127】

さらに、本研究で決定した相関は、上皮核が間質に浸潤するにつれて、上皮核が、典型的には、形状（例えば、変動の少ないさらなる円形）およびサイズ（例えば、領域および境界長）の変動がより小さくなるようになり、色の変動がより小さくなるという仮説によって少なくとも一部を説明することができる。この間質の浸潤はまた、なぜ間質の形態計測的フィーチャが疾患の進行と相関するように決定されたのかを説明することができる。特に、癌画像は、典型的には、少量の間質によって特徴づけられる。これは、癌が進行するにつれて間質領域が上皮細胞の細胞質に置換されるからである。これにより、間質の密度値がより高くなる。これは、間質稠密度が減少し、形状がよりフラクタルになるからである（オブジェクトが変形し、より薄くなるにつれて、オブジェクト半径が領域を超えて増加する）。本研究で決定された相関についてのさらなる根拠は、組織全体を移動する豊富な赤血球が、細胞が前立腺から遊離して外部に播種されるための手段として疾患の進行に関連し得る血管形成または新規の血管形成のいくつかの測定値を反映し得、それにより、PSA/BCR再発の臨床結果に影響を及ぼすことであり得る。

10

【0128】

上記のように、転帰と相関するように本明細書中に提供した教示によって決定された少なくともいくつかの特定の形態計測的フィーチャが、例えば、形態計測的フィーチャを計算するために本発明によって使用される特定のハードウェア、ソフトウェア、またはその組み合わせに依存し得ると理解されるであろう。Definiens Cellengerソフトウェアおよび本明細書中に記載のソフトウェアによって測定された特定の形態計測的フィーチャは例示のみを目的とし、本発明の範囲を逸脱することなく任意の他のハードウェア、ソフトウェア、またはその組み合わせを使用することができる。

20

【0129】

分子分析。IHCによって評価された12のバイオマーカーのうち、全部で43個の固有のフィーチャを記録した（以下の表8a、8b、および8cは、認められたバイオマーカー-分子的フィーチャのまとめを示す）。

30

【0130】

【表 8 - 1】

表8a. 組織学的成分による細胞(%)染色(+)および強度(研究1)

マーカー	腫瘍			PIN			腺		
	1+	2+	3+	1+	2+	3+	1+	2+	3+
Ki-67	23.9 ± 31.38	9.8 ± 21.32	2.4 ± 4.64	25.3 ± 32.50	10.3 ± 21.51	2.6 ± 3.29	1.8 ± 9.96	0.0 ± 0.36	0.1 ± 0.63
中央値	4.7	0.0	0.0	4.8	0.0	0.0	0.0	0.0	0.0
範囲	0.0 - 100.0	0.0 - 100.0	0.0 - 26.3	0.0 - 100.0	0.0 - 100.0	0.0 - 39.5	0.0 - 96.0	0.0 - 4.0	0.0 - 6.3
CK 18	NA	NA	100.0 ± 0.00	NA	NA	100.0 ± 0.00	NA	NA	100.0 ± 0.00
平均 ± SD	NA	NA	100.0	NA	NA	100.0	NA	NA	100.0
中央値	NA	NA	100.0	NA	NA	100.0	NA	NA	100.0
範囲	NA	NA	100.0 - 100.0	NA	NA	100.0 - 100.0	NA	NA	100.0 - 100.0
CD45	NA	NA	0.0 ± 0.04	NA	NA	0.0 ± 0.01	NA	NA	0.0 ± 0.00
平均 ± SD	NA	NA	0.0	NA	NA	0.0	NA	NA	0.0
中央値	NA	NA	0.0	NA	NA	0.0	NA	NA	0.0
範囲	NA	NA	0.0 - 0.4	NA	NA	0.0 - 0.1	NA	NA	0.0 - 0.0
CD68	NA	NA	0.0 ± 0.01	NA	NA	NA	NA	NA	NA
平均 ± SD	NA	NA	0.0	NA	NA	NA	NA	NA	NA
中央値	NA	NA	0.0	NA	NA	NA	NA	NA	NA
範囲	NA	NA	0.0 - 0.1	NA	NA	NA	NA	NA	NA

表8b. 組織学的成分によるCD34細胞(%)染色(+) (研究1)

マーカー	PIN		腫瘍		腫瘍/PIN		PIN/間質	
	間質	腫瘍	間質	腫瘍	間質	腫瘍	間質	
平均 ± SD	0.0 ± 0.05	0.0 ± 0.03	0.1 ± 0.21	0.0 ± 0.06	0.0 ± 0.08	0.0 ± 0.05	0.0 ± 0.05	
中央値	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
範囲	0.0 - 0.4	0.0 - 0.2	0.0 - 0.9	0.0 - 0.5	0.0 - 0.4	0.0 - 0.3	0.0 - 0.3	

【 0 1 3 1】

10

20

30

40

【表 8 - 2】

表8c. 組織学的成分による染色指数(0~300)(研究1)

マーカー	腫瘍	PIN	腺
AR			
平均±SD	179.8±71.4	64.3±75.10	22.6±56.86
中央値	200	36.5	0
範囲	0-300	0-300	0-300
CK14			
平均±SD	2.6±5.83	31.2±57.35	4.7±20.42
中央値	0	0	0
範囲	0-42	0-285	0-150
サイクリンDI			
平均±SD	1.5±5.15	0.0±0.27	0.0±0.0
中央値	0	0	0
範囲	0-33	0-3	0-0
PSA			
平均±SD	128.0±68.85	135.7±97.88	13.9±41.32
中央値	100	111	0
範囲	0-300	0-300	0-201
PSMA			
平均±SD	0.5±2.97	9.5±26.93	2.5±15.00
中央値	0	0	0
範囲	0-21	0-154	0-99
p27			
平均±SD	4.3±9.61	7.0±19.49	2.1±12.03
中央値	0	0	0
範囲	0-80	0-140	0-120
Her-2/neu			
平均±SD	4.1±18.50	0.1±1.00	0.0±0.00
中央値	0	0	0
範囲	0-146	0-10	0-0

10

20

これらの12種の抗体から、PSA再発と相関する14個の分子的フィーチャを含む8つのバイオマーカーを選択した。より高度に選択された分子的フィーチャのいくつかの例を、以下のように注釈づけし(バイオマーカー-モデルによって選択された回数)、これらのフィーチャには、以下が含まれる: AR染色指数-腫瘍(93)、AR染色指数-萎縮性腺(54)、CD34-関連腫瘍/PIN(22)、Ki-67-腫瘍(18)、およびCD45-PINに関連(17)(PINは前立腺上皮内新生物である)。図7aおよび7bは、それぞれARおよびCD34の発現プロフィールを証明する代表的な視野を示す。バイオマーカー発現のプロフィールは、高度に選択された、いくらか異種のARおよびCD34の発現パターンに関して注目に値した。これらのマーカーならびに腫瘍、萎縮性腺(ARについて)、および腫瘍/PIN(CD34について)とのその関係により、PSA再発の臨床転帰に及ぼす生物学および機能的な有意性が示唆される。選択されたマーカーの第2の群はKi-67およびCD45を含み、その両方がARおよびCD34と比較した場合に顕著であるが選択頻度が全体的に低かった。

30

40

【0132】

分析および統計研究。モデルの訓練(N=100)および試験(N=32)のために120個のランダムスプリットを作製した上記の専門知識およびドメイン特異的フィーチャ選択手順の両方を使用して、最終フィーチャセットを全部で41個のフィーチャに削減させた。そのうちの17個は臨床的フィーチャであり、10個は形態計測的フィーチャであり、14個は分子的フィーチャであった。図5は、選択されたフィーチャの完全なリストを示す。10個の形態計測的フィーチャは上記である。臨床的フィーチャおよび分子的フィーチャを、以下にさらに記載する。

【0133】

臨床的フィーチャ

1. 生検グリーンソスコア: 病理学者によって行われた複数の針生検組織サンプルに割り当てるグリーンソスコア分類(優位および二次的)のまとめ。各悪性度の生成によって組織学によって標準化されたいくらか主観的な前立腺癌の構造を示す手段を作製するために、グリーンソスコアリングシステムを構築した。悪性度は、腺単位および上皮細胞の分化度に基

50

づいて1から5までの範囲である。優位（一次）および準優位（二次）パターンを共に加えて、グリーソンサマリー（Gleason Summary）を作成した。さらに、全間質稠密度のフィーチャ、上皮細胞サイズ、および核フィーチャを、時折、全分類システム中で考慮する。

2．人種（例えば、アフリカ系アメリカ人、コーカサスなど）

3．UICC期：International Union against CancerのTNM病期分類システムを、癌の臨床病期分類を定義するために使用する。ここで、「T」は腫瘍サイズを示し、「N」はリンパ節の関与を示し、「M」は遠位部位への転移を示す。

4．倍数性の結果：前立腺癌上皮細胞内の全DNA含量を反映するDNA含量。良性細胞および挙動が良好な（well-behaved）腫瘍細胞は成長し、規則正しく分裂する。静止期では、これらの細胞は1つの完全な染色体組を含む（これは、二倍体条件である）。この完全な染色体組は、Ma由来の23個（すなわちN）の染色体およびPa由来の23個（この場合もN）の染色体（合計で2Nに等しい）からなる。細胞はその染色体数を倍加しなければならず、それにより分裂して2つの完全な染色体組（これは4N、すなわち、四倍体状態である）を作製することができる。分裂の完了後、新規の各細胞は遺伝子材料の半分を受け、再度二倍体（2N）になる。DNA倍数性分析がこれらの細胞群に対して行われる場合、ほとんどの細胞が二倍体であり、その小画分（分裂する準備を進めている細胞）が四倍体であろう。さらに、各細胞中の遺伝子材料の量の測定およびグラフの作成では、主な二倍体ピークおよび小さな四倍体ピークが認められるであろう。細胞中のDNA量を、遺伝子材料に結合する色素でのDNAの染色によって測定することができる。この色素（Feulgen stain）の濃度および分布を、画像分析マイクロアレイによって測定することができる。

【0134】

腫瘍が悪化した場合、腫瘍は、かつて分裂していたように規則的には分裂しない傾向がある。完全な染色体組を有する静止状態の代わりに、静止期は1つの組または半分のみを有し得る。かかる細胞は、二倍体および四倍体のいずれでもなく、その中間であるDNA含量を有するであろう。上記グラフ上のこれらの細胞のプロットにより、他の2つのピークの間の変数性ピークが得られるであろう。研究により、有意な変数性ピークを有する腫瘍は、ピークを持たない腫瘍と同様に、挙動しないことが示された。倍数性状態と核グレードとの間に強い相関が存在するので、これは驚くべきことではない。核グレードを、十分な前立腺癌の経験を有する任意の病理学者によって評価することができる。DNA倍数性分析が付加される価値は、標準化された技術を使用して研究所間で比較することができる、グリーソンスコアリングの近似精度を迅速にチェックするために使用することができる客観的測定であることである。例えば、変数性ピークを有する任意のグリーソンスコア $2 + 2 = 4$ または $2 + 3 = 5$ 個の腫瘍を、スコアの予想される調整のために、スコアを潜在的に再評価すべきである。

5．DREの結果：触診による前立腺内および前立腺外の拡大の範囲を決定するために使用される直腸の指診の結果（例えば、陰性または陽性）。

6．リンパ節の関与：臨牀的/外科的検査または前立腺摘除時のいずれかによって評価することができるリンパ節が腫瘍細胞（例えば、前立腺癌上皮細胞）を含む範囲の測定。

7．優位生検グリーソン分類：上記の生検グリーソンスコアの説明を参照のこと。これは、生検または前立腺摘除検体で認められた優位グリーソン分類パターンを反映する。

8．S期の倍数性率：細胞周期の増殖期またはS期にあり、且つ腫瘍の潜在的成長を反映する細胞内容物の割合を示す。

9．術後グリーソンスコア：前立腺切除術サンプルの種々の領域由来の術後に採取した組織のスコアリング。

10．TNM期：前立腺摘除後のUICCおよび組織サンプルの病理学的試験に基づいた腫瘍、結節、および転移。

11．優位術後グリーソン分類：前立腺摘除検体中に存在する最も支配的な組織学的フィ

10

20

30

40

50

ーチャを示す優位グリーソン分類。

12. 年齢
13. 精囊の関与：腫瘍による精囊の浸潤。
14. 術前PSA：手術前に認められるPSAレベル。
15. 倍数性画分比率：上記の倍数性の結果を参照のこと。
16. 切除縁の関与：腫瘍/前立腺が手術時に除去された床が腫瘍細胞を含む範囲を反映する腫瘍による切除縁の関与。
17. 囊外の関与：前立腺被膜を超えた腫瘍の拡大。

【0135】

分子的フィーチャ

1. AR-腫瘍：腫瘍についてのアンドロゲン受容体(AR)染色指数。これは、ARについて陽性染色された細胞の比率および強度の基準である。前立腺癌に関して、染色指数は、評価した前立腺サンプル中の上皮細胞の核中で検出される褐色の反応生成物の程度を示すことができる。

2. AR-腺：腺構造内に存在する腫瘍についてのAR染色指数。

3. CD34-腫瘍/PIN：腫瘍およびPINに関連する血管の内皮細胞へのCD34の局在化。

4. Ki67-腫瘍2：腫瘍上皮細胞核内のKi67陽性核の同定。

5. CD45-PIN3：PINに関連するCD45陽性リンパ球の同定。

6. CD34-腫瘍/間質：腫瘍に関連する血管へのCD34の局在化。

7. Ki-67-腫瘍3：上記を参照のこと。

8. p27-腫瘍：腫瘍上皮細胞の核内のp27の同定。

9. C14-PIN：腺単位の(上皮)基底細胞中のサイトケラチン14の同定。

10. CD34-腫瘍：腫瘍に関連する血管に対するCD34の局在化。

11. PSA-腺：腺単位の管腔上皮細胞に対するPSAの同定。

12. PSMA-PIN：PINとして同定された領域の腺/管腔組織に対するPSMAの同定。

13. CD34-PIN/間質：PINに関連する血管へのCD34の局在化。

14. CD45-腫瘍3：腫瘍に関連するCD45陽性リンパ球の同定。

【0136】

SVRCを使用してこの過程にデータの各ドメインを分析した場合、モデルの予測精度が増加した。臨床データのみを調査する場合、内部検証を使用して、一致指数は0.79であった。分子ドメイン由来のフィーチャの付加により、一致指数は0.81に増加した。形態計測的フィーチャの付加によって形成された最終モデルの一致指数は、0.84に到達した。これらの内部検証された各モデルも外部検証し(上の材料と方法に記載のように)、同一の傾向が認められた。最終的に選択されたフィーチャセットに対してNNciを使用して、一致指数は0.88に到達した。

【0137】

NNciおよびSVRCモデルの得られた出力を、各患者についてのPSA再発の相対リスクの評価と解釈することができる。このスコアの四分位値(25%、>25%~75%、>75%)を使用して、患者のリスク群を作製した；NNciモデルに従った各リスクについての再発のカプラン・マイヤー評価を図8に示す。群は、PSA再発までの期間に統計的有意差を示した(ログランク検定、 p 値<0.0001)。 p 値は、機会のみが、層(これらの実施例中のリスク群)の間の認められた差異を担う確率を示す。したがって、 p 値がより低いと真の統計的関連がある可能性がより高くなる。一般に、0.05以下の任意の p 値は統計的に有意である。

【0138】

研究2。本コホート中の患者268人(研究1で分析した患者132人のうちの129人を含む)について、診断時の年齢の中央値は63歳(最年少：38歳、最年長81歳)であり、根治的前立腺摘除前PSAの中央値は7.8ng/dl(最小値：0.9、最大

10

20

30

40

50

値：81.9)であった。前立腺摘除サンプルに基づいて、腫瘍の40%はグリーソンスコアが7未満である一方で、前立腺摘除術の55%がグリーソン7であった。前立腺摘除術の残りの5%は、グリーソンスコアが7超であった。患者157人(59%)は、pT2N0M0疾患と診断され、患者72人(27%)はpT3aN0M0と診断され、残りの患者39人(14%)はpT3bN0M0またはpT1-3N+と診断された(このコホートについての分析された全臨床病理学的フィーチャの詳細については、前出の表5を参照のこと)。38人(14%)の患者はPSA再発を経験した一方で、残りの患者(86%)は打ち切った。打ち切られた患者について、追跡期間の中央値は58.7ヶ月(すなわち、5年を少し下回る)であった。PSA再発までの全時間の中央値に届かなかった。PSA再発が予測される以下の3つの臨床的フィーチャを選択した：TNM 臨床病期、切除縁、およびリンパ節。

【0139】

画像分析および形態計測研究。画像分析ソフトウェアのアップデートバージョンを使用するが同一のH&E染色スライドを分析して、全部で350個の形態計測的フィーチャを生成した(上記の表2に示す)。

【0140】

図9は、350個のフィーチャのうちのPSA再発が予測される6個の形態計測的フィーチャを選択したことを示し、これらの形態計測的フィーチャは、上皮核、間質、細胞質、赤血球、および管腔の病理学的オブジェクトに関連した(すなわち、EpithelialNucleiMinCompactness0215、StromaMaxStdDevChannel30569、CytoplasmStdDevMaxDiff0148、RedBloodCellMeanAreaPx10386、RedBloodCellStdDevAreaPx10388、およびLumenMinAsymmetry0295)。より詳細には、本研究では、転帰と関連する上皮核の稠密度、間質の青色チャネル値、細胞質の最大差(例えば、細胞質の全色チャネルにわたるその最大値から引いた細胞質に属する最小平均値。結果をオブジェクト輝度で除する)、赤血球領域、および管腔の非対象性の形態計測的フィーチャを選択した。

【0141】

これらの関連のうちの少なくともいくつかについての種々の可能な理由を、上記の実施例1および/または研究1と併せて記載する。例えば、形態計測的フィーチャ(上皮細胞の稠密度、浸潤上皮細胞による間質の変動および破壊、および管腔サイズ減少の証拠が含まれる)は全て、より高いグリーソン分類(すなわち、より高い疾患の進行)の組織学的証拠を提供するであろう。より高いグリーソン分類により、より浸潤性の高い前立腺腫瘍が示唆され、これは、術後PSA再発を支持する腫瘍の転移および/または拡大を支持するであろう。さらに、種々の形式の赤血球の同定により、多数の血管が示唆されるであろう。さらなる血管の証拠により、上皮細胞が前立腺から抜け出し、PSAを産生する外部に分布することができる可能な経路が作製されるであろう。

【0142】

研究2で選択された臨床的フィーチャおよび分子的フィーチャを図9に示し、以下に列挙する。これらの臨床的フィーチャおよび分子的フィーチャは、上記で説明されている。

【0143】

臨床的フィーチャ

1. TNM期
2. 切除縁の関与
3. リンパ節の関与

分子的フィーチャ

1. AR 染色指数(腫瘍)

図9中の各数字は、対応するフィーチャに基づいた予測モデルの一致指数を示し、図9中の全ての他のフィーチャは、より小さな数字を有する。例えば、0.8483は、フィーチャ(TNM臨床病期、切除縁、EpithelialNucleiMinComp

10

20

30

40

50

ctne0215、リンパ節、およびStromaMaxStddevChannel30569)に基づいたモデルのCIである。同一の5個のフィーチャ+AR染色指数(腫瘍)に基づいたモデルのCIは0.8528である。言い換えれば、モデルへのAR染色指数分子的フィーチャの付加により、モデルの予測力が増加する。

【0144】

分子分析。さらなる免疫組織化学研究は必要なかった。最初に収集したデータを、材料と方法に記載のように使用した(分子的フィーチャの完全なまとめに付いては、表9a、9b、および9cを参照のこと)。

【0145】

【表 9 - 1】

表9a. 組織学的成分による細胞(+)および強度(研究2および研究3)

マーカー	腫瘍			PIN			腺		
	1+	2+	3+	1+	2+	3+	1+	2+	3+
Ki-67	22.1 ± 30.30	7.3 ± 17.04	1.9 ± 4.01	23.2 ± 31.36	7.9 ± 18.16	2.0 ± 4.46	1.3 ± 7.96	1.2 ± 9.78	0.3 ± 1.55
平均±SD	1.3	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
中央値	0.0 - 100.0	0.0 - 100.0	0.0 - 26.3	0.0 - 100.0	0.0 - 100.0	0.0 - 39.5	0.0 - 96.0	0.0 - 96.5	0.0 - 13.0
範囲									
CK 18	NA	NA	100.0 ± 0.00	NA	NA	1.0 ± 0.04	NA	NA	100.0 ± 0.00
平均±SD	NA	NA	100.0	NA	NA	100.0	NA	NA	100.0
中央値	NA	NA	100.0 - 100.0	NA	NA	0.5 - 100.0	NA	NA	100.0 - 100.0
範囲									
CD45	NA	NA	0.0 ± 0.04	NA	NA	0.0 ± 0.01	NA	NA	0.0 ± 0.00
平均±SD	NA	NA	0.0	NA	NA	0.0	NA	NA	0.0
中央値	NA	NA	0.0 - 0.4	NA	NA	0.0 - 0.1	NA	NA	0.0 - 0.0
範囲									
CD68	NA	NA	0.0 ± 0.01	NA	NA	NA	NA	NA	NA
平均±SD	NA	NA	0.0	NA	NA	NA	NA	NA	NA
中央値	NA	NA	0.0 - 0.1	NA	NA	NA	NA	NA	NA
範囲									

表9b. 組織学的成分によるCD34細胞(%)染色(+) (研究2および研究3)

マーカー	腫瘍		PIN		腺	
	間質	腫瘍	間質	腫瘍	間質	腫瘍
平均±SD	0.0 ± 0.04	0.0 ± 0.11	0.1 ± 0.18	0.0 ± 0.08	0.0 ± 0.08	0.0 ± 0.04
中央値	0.0	0.0	0.0	0.0	0.0	0.0
範囲	0.0 - 0.4	0.0 - 1.7	0.0 - 0.9	0.0 - 0.6	0.0 - 0.4	0.0 - 0.3

【 0 1 4 6 】

【表 9 - 2】

表9c. 組織学的成分による染色指数(0~300)(研究2および研究3)

マーカー	腫瘍	PIN	腺
AR			
平均±SD	172.1 ± 75.3	79.6 ± 82.74	28.9 ± 67.25
中央値	200	66.0	0
範囲	0 - 300	0 - 300	0 - 300
CK14			
平均±SD	2.1 ± 6.32	34.4 ± 61.46	8.5 ± 32.62
中央値	0	0	0
範囲	0 - 69	0 - 300	0 - 300
サイクリンD1			
平均±SD	1.4 ± 6.99	0.0 ± 0.21	0.0 ± 0.0
中央値	0	0	0
範囲	0 - 90	0 - 3	0 - 0
PSA			
平均±SD	118.3 ± 71.10	139.4 ± 97.16	22.8 ± 55.14
中央値	100	134	0
範囲	0 - 300	0 - 300	0 - 300
PSMA			
平均±SD	0.2 ± 2.09	6.4 ± 21.02	2.9 ± 22.94
中央値	0	0	0
範囲	0 - 21	0 - 154	0 - 300
p27			
平均±SD	3.9 ± 8.20	6.4 ± 18.83	1.3 ± 8.65
中央値	0	0	0
範囲	0 - 48	0 - 140	0 - 120
Her-2/neu			
平均±SD	3.4 ± 16.69	0.2 ± 1.12	0.0 ± 0.00
中央値	0	0	0
範囲	0 - 150	0 - 10	0 - 0

10

20

PSA再発が予測される以下の1つの分子的フィーチャを選択した：AR 染色指数 - 腫瘍。

【0147】

分析および統計研究。専門知識および簡潔なブーストラッピングを使用して、方法は、一致指数(CI)0.87の10個のフィーチャ(3つの臨床病理学的フィーチャ、6つの形態計測的フィーチャ、および1つの分子的フィーチャ)のサブセットを見出した(図9は、選択されたフィーチャの完全なリストを示す)。得られたSVRCモデルの出力も、各患者についてのPSA再発の相対リスクの評価と解釈することができる。このスコアの四分位値(<25%、>25%~75%、>75%)を使用して、患者のリスク群を製作した；SVRCモデルによって予想される各リスク群についての再発のカプラン・マイヤー評価を図10に示す。群は、PSA再発までの期間に統計的有意差を示した(ログランク検定、p値<0.0001)。

30

【0148】

研究3。本研究は、患者の臨床病理学的特徴が同一であるように、研究2と同一のコホートを使用した。転帰に関して、19人(7%)の患者が任意の理由によって死亡した一方で、残りの患者(93%)はその最後の来院で生存しており、データ収集を打ち切った。打ち切られた患者について、追跡期間の中央値は64.8ヶ月(すなわち、5年を少し超過する)であった。死亡までの全時間の中央値に届かなかった。任意の原因による死亡が予測される以下の2つの臨床的フィーチャを選択した：TNM臨床病期および患者の年齢。

40

【0149】

画像分析および形態計測研究。研究2由来の同一の350個の形態計測的フィーチャセットを、本研究で使用した。図11は、350個のフィーチャのうち、任意の原因による死亡が予測される11個の形態計測的フィーチャを選択し、ここで、これらのフィーチャは、間質、赤血球、および上皮核の病理学的オブジェクト(すなわち、StromaMinMeanChannel10535、RedBloodCellMeanStdDev

50

Chann30474、StromaMinMeanChannel20539、RedBloodCellMinMeanChannel20443、RedBloodCellStddeStddeChann20472、StromaMaxMaxDiff0529、EpithelNucleMeanBordeLengtPx10206、EpithelialNucleiMeanAreaPx10194、EpithelNucleiStddevElliptFit0228、RedBloodCellStddeStddeChann30476、およびRedBloodCellStddevElliptiFit0420、ここで、「チャンネル」は、赤色(R)、緑色(G)、および青色(B)の画像色チャンネルをいう)に関連したことを示す。より詳細には、本研究では、赤色チャンネルの平均値、青色チャンネルの平均値、および間質の最大差の形態計測的フィーチャを、転帰と相関するように決定した。赤色チャンネルの平均および標準偏差、緑色チャンネルの平均および標準偏差、および赤血球の楕円適合度の形態計測的フィーチャを、転帰と相関するように決定した。楕円適合度の形態計測的フィーチャを決定するために、赤血球と同一領域を有する楕円を作製し、楕円の外側の赤血球の領域を、赤血球で満たされていない楕円の内側の領域と比較し、適合しなかった場所に値0を割り当てたのに対して、完全に適合したオブジェクトに値1を割り当てた。上皮核の境界長、領域、および楕円適合度の形態計測的フィーチャを、転帰と相関するように決定した。

10

【0150】

これらの相関のうち少なくともいくつかについての種々の可能な理由は、実施例1および/または研究1と併せて上に記載されている。例えば、上皮核の全形状は、より高いグリーソン分類の組織学的外観を反映する。さらに、本研究では、間質に関連する相関を、癌がその上皮細胞の妨害によって進行するにつれて間質のコントラストが減少する(最大に異なる形態計測的フィーチャによって測定)という理解によって説明することができる。

20

【0151】

分子分析。研究2由来の同一の分子的フィーチャセットを、本研究で使用した。任意の原因による死亡が予想される以下の1つのフィーチャを選択した：PSA 染色指数 - 萎縮性腺。

【0152】

分析および統計研究。本コホートでは、全部で14個のフィーチャ(2個の臨床病理学的フィーチャ、11個の形態計測的フィーチャ、および1個の分子的フィーチャ)を選択した。最終モデルの一致指数(CI)は0.80であった。選択されたフィーチャの完全なリストを、図11に示し、以下に列挙する。選択された臨床的フィーチャおよび分子的フィーチャを以下に列挙する。臨床的フィーチャを上記で説明する。

30

【0153】

臨床的フィーチャ

1. TNM期
2. 年齢

分子的フィーチャ

1. psapsi: 前立腺上皮内新生物(PIN)中の前立腺特異抗原(PSA)の染色指数をいう。

40

【0154】

図11中の各数字は、対応するフィーチャに基づいた予測モデルの一致指数を示し、図11中の全ての他のフィーチャは、より小さな数字を有する。例えば、0.6804は、StromaMinMeanChannel10535に基づいたモデルのCIであり、0.7362は、モデルがStromaMinMeanChannel10535およびTNMの両方に基づいた場合のCIである。

【0155】

得られたSVRcモデルの出力も、各患者の死亡の相対リスクの評価と解釈することができる。このスコアの四分位値(<25%、>25%~75%、>75%)を使用して、

50

患者のリスク群を作製した。SVRCモデルによって予想される各リスク群についての再発のカプラン・マイヤー評価を図12に示す。ログランク検定を使用して、リスク群間で生存の有意差が認められた ($p < 0.0001$)。

【0156】

結果の考察 (実施例2)

モデルの予測精度を保持しながら研究1(41)から研究2(10)までの選択されたフィーチャ(複合)の減少が認められたことにより、異なる機械学習法によって達成された精度およびフィルタリング属性を強調された。患者268人のコホートで構築されたモデルの一致指数は0.87であった。比較すると、Kattanノモグラム(20)をこのコホートに適用した場合、一致指数0.78が達成された。おそらく、研究2で考察された上記モデルが感度80%で早期PSA再発(5年以内)する患者を正確に分類する能力がより顕著である。比較すると、Kattanノモグラムは、たった54%の感度で同一の予測を行うことができる。これは、かかる予測試験が早期介入の決定で機能を果たすという役割をさらに強調した。最後に、示したモデルの出力を使用して、時間枠内に関して任意の表示を行わずに所与の年数内で患者が再発する確率を1度評価するのと対照的に、長期にわたって患者が再発する可能性を評価することができる。

10

【0157】

研究3では、研究2由来の既存の専門知識を使用し、全生存のための予測モデルを構築することが目的であった。首尾の良い最終結果は、全部で14個の組み合わせたドメインフィーチャを使用して、精度80%で個体の全生存および死亡までの期間を予測することができることであった。少数の事象(任意の原因による7%の死亡)および比較可能な公開されたノモグラムによって制限されるにもかかわらず、結果は、これらの予測試験型の構築のためのシステムアプローチの使用をさらに支持する。

20

【0158】

独立した外部検証研究を用いる遡及的多施設集団を使用した不良転帰(すなわち、前立腺癌に起因する転移および/または死亡)の臨床的測定を含めるための全患者の分析の拡大に関してさらに努力されている最中である。さらに、手術前の治療上の問題点に影響を及ぼすために診断的針生検を調査するための「システム病理学」アプローチが最近開始されている。

【0159】

上記実施例は、臨床的フィーチャ、腫瘍組織形態計測、および分子分析を組み合わせる「システム病理学」プラットフォームが首尾よく構築されたことを証明している。専門知識および打ち切りデータのサポートベクトル回帰(SVRC)の使用により、フィーチャを3つのドメインから選択し、これを使用して、PSA再発および全生存の予測モデルを構築した。腫瘍診断、患者の予後診断、および特定の療法に対する反応の予測ツールに関連するので、この新規の「システム病理学」アプローチが個別化医療分野で広く適用されていると理解されるであろう。

30

【0160】

(実施例3)

前立腺摘除後の浸潤性疾患の予測

臨床データおよび形態計測データ

前立腺摘除した患者がその後起こす浸潤性疾患(すなわち、前立腺癌の骨への転移を示す陽性骨スキャンによって証明される臨床的な失敗)を予測するために、本研究に着手した。本発明の前に、かかる予測を行うための正確な分析ツールは存在しなかった。上記のように、本発明のシステム病理学アプローチは、PSA再発を正確に予測することを示した。本研究は、本発明を使用して前立腺摘除後の遠位骨転移を正確に予想することもできることを証明する。

40

【0161】

根治的前立腺摘除を受けた患者119人のコホートに、前立腺摘除検体から構築した組織マイクロアレイ(TMA)を組み込んだ研究を行った。ヘマトキシリンおよびエオシン

50

(H & E) 染色組織切片を使用して形態計測 (すなわち、画像分析) 研究を行い、生物学的決定要因を前立腺癌の進行についてのその潜在的な生物学的関連性について選択した一連のバイオマーカーを使用した免疫組織化学 (IHC) を使用して評価した。臨床的な失敗 (すなわち、陽性骨スキャン) のための予測モデルは、管理された多変量学習を介して選択されたフィーチャセットに由来した。各ドメイン中に完全な非欠測データを有する患者 (n = 116) を、打ち切りデータ (SVRC) を取り扱うために開発された回帰のためのサポートベクトルマシンを使用して評価した。モデルの予測性能を、リスク群を定義するために使用した生成スコアを用いた一致指数 (CI) を使用して評価した。

【0162】

患者116人から、患者61人のサブセットを、その臨床的フィーチャに基づいて選択し、このサブセットは、骨転移によって同定された臨床的に失敗した20人の個体を含む。このコホートを使用して、前立腺摘除から5年以内に陽性骨スキャンの可能性を予測するためのモデルを作製した。精度89%ならびに感度および特異性がそれぞれ86%および85%で臨床的な失敗を予測した図13中に示した7個のフィーチャ(4個の臨床的フィーチャおよび3個の形態計測的フィーチャが含まれる)を選択した。選択された形態計測的フィーチャは、細胞質および管腔の病理学的オブジェクトと関連していた。より詳細には、選択された形態計測的フィーチャは、総組織領域で除した細胞質領域、総組織領域で除した管腔領域、および平均赤色チャネルの細胞質標準偏差であった。臨床的フィーチャを、以下に列挙する。

10

【0163】

臨床的フィーチャ

1. 囊外拡大 (ECE)
2. 精嚢浸潤 (SVI)
3. 優位前立腺摘除グリーソン分類 (PGG1)
4. リンパ節浸潤 (LNI)

考察

臨床的フィーチャの形態計測的フィーチャとの組み合わせにより、前立腺摘除後5年以内の臨床的な失敗を予測するための第1の正確な予後試験が得られた。記載のように、この試験は、患者が前立腺摘除後5年以内に臨床的な失敗をする可能性が最も高いこと(および時期)を精度89%で予測することができる。モデルの臨床的フィーチャおよび形態計測的フィーチャへの分子的フィーチャの付加の結果は、現在、調査中である。

20

30

【0164】

(実施例4)

肝臓毒物学

形態計測データ

毒物学分野において画像分析および統計的モデル化能力を証明するために、本研究に着手した。具体的には、研究は、ラット肝臓切片の獲得および分析を必要とし、全ての目的は、切片を正常または異常に分類することである。この過程を自動化すると同時に高レベルの分類精度を達成することができることにより、予測研究において毒性を客観的にスクリーニングするための高処理プラットフォームを作製することができる。

40

【0165】

研究を以下の2つの段階に分けた。初期に、訓練組として100個のラット肝臓切片組(80個の正常肝臓切片および20個の異常な肝臓切片)を使用した。この切片組を使用し、上記組織画像分析システムを使用して画像分析アプリケーションを開発し、フィーチャおよびモデルを選択して切片を分類した。次いで、確立された画像分析過程を、第2期で100個の非標識ラット肝臓切片組に適用し、訓練期でデザインされた統計的モデルを試験した。

【0166】

セグメント化精度

病理学者の評価によって測定された全オブジェクトの総セグメント化精度は、80% ~

50

90%であった。

【0167】

統計

研究の統計成分 (statistical component) は、2つの工程を含んでいた。第1の工程は、切片の画像分析によって生成された画像化データからのフィーチャの選択を含んでいた。分類のためのフィーチャの使用数の減少により、切片分類のロバスト性および信頼性を改善することができる。第2の工程は、各切片 (異常、正常) のために選択されたフィーチャセットおよび標識を使用したモデルの訓練および標識が未知である独立したラット肝臓切片組の分類の予測による試験の両方を含んでいた。

【0168】

フィーチャの選択

上記の各オブジェクトのために生成した統計的測定は以下であった。

【0169】

- オブジェクト数
- 相対領域 (総画像領域に対する比率)
- 最小サイズ (ピクセル)
- 最大サイズ (ピクセル)
- 平均サイズ (ピクセル)
- サイズの標準偏差

切片あたり複数の画像を分析したので、これらの測定自体を、各ラット肝臓切片の全画像から平均した。元のフィーチャの総数は378であった。

【0170】

フィーチャ選択も2つの工程を含んでいた。第1の工程は、専門知識を使用した。病理学者は、切片の画像分析によって生成された元のフィーチャリストからフィーチャを選択した。フィーチャを含めるか排除するかどうかの決定は、肝臓病理学の理解および遭遇し得る潜在的な異常性/毒性に基づいた。専門知識を使用して、378個の元のフィーチャセットから、90個のフィーチャを選択した。

【0171】

次いで、これらのフィーチャを段階的判別分析を使用して試験して、分析のためのフィーチャ数をさらに減少させた。各クラスを作り上げたフィーチャセットは、一般的な共分散行列を使用した多変量正規であると考えられる。共分散分析からF検定の有意レベルにしたがってモデルを出入りさせるようにフィーチャを選択した。ここで、すでに選択したフィーチャは共変量として作用し、検討中のフィーチャは従属変数である。有意レベル0.15を使用した。

【0172】

- モデル中でフィーチャから段階的選択を開始しなかった。各工程では、モデルを試験した。

【0173】

- ウィルクスの (尤度基準) によって測定されたモデルの識別力への寄与が最小であるモデル中のフィーチャがとどまるための基準を満たすことができなかった場合、そのフィーチャを除去した。

【0174】

- そうでなければ、モデルの識別力に最も寄与したモデル中に入らないフィーチャを入れた。

【0175】

- モデル中の全フィーチャがとどまるための基準を満たし、他のフィーチャが入るための基準を満たさなかった場合、段階的選択過程を停止した。

【0176】

分類 / モデルの訓練

次いで、選択されたフィーチャを線形判別分析 (LDA) に入力し、肝臓切片を異常ま

10

20

30

40

50

たは正常に分類した。モデルの出力を、クロスバリデーションによって潜在的バイアスに対して補正した。

【0177】

神経回路網も分類子として調査した。選択されたフィーチャを、神経回路網モデルへの入力として使用した。このモデルは、ゼロ隠れユニット (zero hidden unit) および入力層と出力層との間の直接接続を有する標準的な多層パーセプトロン (MLP) 構造である。モデルを、以下に説明するROC曲線下領域への近似を直接最大化するように試みることによって訓練した。この基準によって訓練されたMLPモデルは、典型的な基準 (例えば、平均平方誤差およびクロスエントロピー) によって訓練されたMLPよりも高い精度が達成されることが見出された。

10

【0178】

両モデル由来の出力を使用して、モデル出力の異なる値を切点として選択し、各切点について感度および特性を計算し、これらを二次元プロット (y軸に沿った感度およびx軸に沿った特異性) 中にプロットすることによって受信者動作特性 (ROC) 曲線を作製した。ROC曲線下領域 (AUC) は、各モデルの精度を評価するために両方の基準を使用し、モデルが肝臓切片を異常または正常と正確に分類する能力と解釈することができる。典型的には、感度および特異性を、それぞれ、真の陽性率および陰性率に関して記載する。したがって、本研究の文脈において、異常クラスを「陽性」結果と見なした一方で、正常クラスを「陰性」クラスと見なした。したがって、感度は、真の陽性率 (すなわち、正確に異常に分類された肝臓切片の比率) であり、他方では、特異性は、真の陰性率 (すなわち、正確に正常に分類された肝臓切片の比率) である。

20

【0179】

ROC曲線から、訓練組由来の選択された感度および特異性を、以下の結果の項に示す。

【0180】

モデル試験

一旦開発されると、線形判別関数および神経回路網の両方のパラメータをロックした。ラット肝臓画像の試験組からの統計学的測定値の受信の際、各モデル出力のクロスバリデーションの結果をそれぞれ使用して評価した各切点を使用して、両分類子を適用した。切点は共に将来の産業規模の適用のための感度100%および特異性90% (共にクロスバリデーションに基づく) に対応した。肝臓のこの外部検証組の最初の評価のために、モデルの精度の評価を、肝臓切片の真の分類に対して非盲検の第三者によって行った。次いで、この第三者により、結果を検証するための試験キー (test key) が提供される。

30

【0181】

結果

両モデルについてのROC曲線下領域は1に非常に近く、異常および正常な肝臓切片のほぼ完全な区別を示す。LDAを使用して導いた関数のAUCは0.99であり、神経回路網を使用して導いた関数のAUCは0.98である。

【0182】

ROC曲線において、各モデルの感度および特異性も認められ、これは、肝臓切片を異常または正常に分類するためのモデル出力に適用した切片に依存した。表10は、感度 - 特異性対の選択をまとめている。

40

【0183】

【表 1 0】

LDA		NN	
特異性	感度	特異性	感度
100%	65%	100%	65%
99%	75%	99%	70%
98%	100%	98%	85%

10

表10

試験

試験キー標識を、線形判別関数および神経回路網の予測される分類と比較した。キーに基づいて、以下のように結果を図 1 1 a および図 1 1 b にまとめる。

【 0 1 8 4】

【表 1 1 - 1】

試験キー標識

		試験キー標識		
		異常	正常	
LDA 標識	異常	42 (TP)	19 (FP)	100
	正常	7 (FN)	32 (TN)	
		49	51	

20

$$\text{感度} = \text{TP}/(\text{TP}+\text{FN}) \times 100 = 42/(42+7) \times 100 = (42/49) \times 100 = \mathbf{86\%}$$

$$\text{特異性} = \text{TN}/(\text{FP}+\text{TN}) \times 100 = 32/(19+32) \times 100 = (32/51) \times 100 = \mathbf{63\%}$$

30

表11a

【 0 1 8 5】

【表 1 1 - 2】

試験キー標識

		試験キー標識		
		異常	正常	
NN 標識	異常	36 (TP)	19 (FP)	100
	正常	13 (FN)	32 (TN)	
		49	51	

40

$$\text{感度} = \text{TP}/(\text{TP}+\text{FN}) \times 100 = 36/(36+13) \times 100 = (36/49) \times 100 = \mathbf{73\%}$$

$$\text{特異性} = \text{TN}/(\text{FP}+\text{TN}) \times 100 = 32/(19+32) \times 100 = (32/51) \times 100 = \mathbf{63\%}$$

表11b

LDA 分類子のために使用した切点は 0 . 0 0 3 1 であった。NN 分類子のために使用 50

した切点は0.0002であった。共に感度100%および特異性90%のシステム要件に対応している。

【0186】

考察

試験セットへの適用後の各分類子の感度および特異性に基づいて、LDAはNNより優れていた。LDA分類子は、感度86%を達成した。これは、感度73%の神経回路網分類子とは対照的に、この分類子が異常なラット肝臓切片をその時に86%の異常として正確に標識したことを意味する。両分類子の特異性は63%であった。各モデルの感度および特異性の両方は、以前に認められた値より低い、任意の分類子の外部セットに対する一般化によってしばしばその精度が低下するので、これは驚くべきことではない。本研究は、画像化技術および統計的モデル化技術の首尾の良い適用を証明した。

10

【0187】

(実施例5)

前立腺癌再発の予測

臨床データ、形態計測データ、および分子データ

別の研究を行い、根治的前立腺摘除を受けた患者の前立腺癌の再発までの期間を予測するモデルを生成した。実施例2と同様に、再発までの期間を、根治的前立腺摘除からPSA(生化学的)再発までの期間(月)と定義した。本研究で使用した前立腺組織の切片は、主に腫瘍から構成されるが、良性エレメントも含んでいた。

20

【0188】

本研究は、実施例2で評価した同一の17個の臨床的フィーチャ(表4)および43個の分子的フィーチャ(表6)についての情報に基づいた。496個の形態計測的フィーチャセット(表1)を、例えば、前立腺癌再発分野の専門知識およびさらなる実験に基づいて、表17(本明細書に添付)中に示す38個のフィーチャに減少させた。実施例2と併せた上記の患者539人のコホート由来の患者262人についての臨床情報、分子情報、および形態情報を、本研究で評価した。形態計測的フィーチャ数が496個から38個に減少したフィルタリング以外に、本研究と実施例2との間の主な相違は、本研究がフィーチャ選択のためのSVRCフィーチャ減少法を使用したことである。SVRCフィーチャ減少は、2006年5月22日に出願された共同所有の米国特許出願第11/438,789号(その全体が本明細書中で参考として援用される)に記載されている。

30

【0189】

結果

6個のフィーチャ(3個の臨床的-病理学的フィーチャ、1個の分子的フィーチャ、および2個の形態計測学的フィーチャ)に基づき、一致指数(CI)0.83の最終モデルを、研究結果として生成した。モデルに含まれる6個のフィーチャを、最終モデルに対するその各フィーチャの寄与と共に図14に示す。PSA再発が予測されるように選択された3個の臨床的フィーチャは、精囊の関与(フィーチャ寄与=-5.2103)、切除縁の関与(-7.3159)、およびリンパ節の関与(-9.3742)であった。選択された1つの分子的フィーチャは、腫瘍中に存在するアンドロゲン受容体(AR)染色指数(-3.5404)であった。これらの臨床的フィーチャおよび分子的フィーチャは、実施例2と共に上に記載されている。選択された2個の形態計測的フィーチャは、総組織領域で除した上皮細胞によって占められる領域(3.2975)および総組織領域で除した間質によって占められる領域(-.34225)であった。例えば、総組織領域は、細胞質、上皮核、管腔、赤血球、間質核、間質、および人工物の領域の合計(1920000ピクセルが最大であり、ここで、画像サイズは1200x1600ピクセルである)を含むことができる。上皮核および間質細胞に関連する形態計測的フィーチャの選択についての可能な理由も、検証研究と併せて以下に記載している。

40

【0190】

5年前の前立腺癌再発を正確に予測するために、最終訓練モデルは、感度82%および特異性81%であった。得られたSVRCモデルの出力を、各患者についてのPSA再発の

50

相対リスク評価と解釈することもできる。このスコアの四分位値 (< 25%、> 25% ~ 75%、> 75%) を使用して患者のリスク群を作製し、各リスク群について再発のカプラン・マイヤー評価を行った。群は、PSA再発までの期間において統計的有意差を示した(ログランク検定、p値 < 0.0001)。

【0191】

検証研究

患者61人からなる外部コホートを使用して最終モデルを検証した。最初の5年以内に前立腺癌再発を経験するリスクのある患者の同定を検証するために、最終モデルは、CI 0.80、感度91%、および特異性70%であった。

【0192】

最終モデルで選択された2つの形態計測的フィーチャの有意性をさらに理解するために、各形態計測的フィーチャについてのカプラン・マイヤー曲線を作成した。漸増量の間質(p = 0.004)および上皮核(しかし、統計的に有意でない、p = 0.28)が好ましい結果と独立して関連することが認められた。これにより、これらの画像パターン由来の定量的測定がグリーンソン分類システムのより客観的な決定要因を示すことができる可能性が高まった。

【0193】

本研究は、臨床的に有用な予測試験を作製するのに、限られた臨床的フィーチャ、分子的フィーチャ、および形態計測的フィーチャの組み合わせが必要ないことを表明した。モデルの予測精度も保持しながらこのフィーチャを減少させた。

【0194】

(実施例6)

前立腺癌再発の予測

臨床データ、形態計測データ、および分子データ

さらに別の研究を行って、根治的前立腺摘除を受けた患者における前立腺癌再発までの期間を予測するモデルを生成した。実施例2および5の場合、再発までの期間を、根治的前立腺摘除からPSA(生化学的)再発までの期間(月)と定義した。本研究で使用した前立腺組織の切片は、主に腫瘍から構成されるが、良性エレメントも含んでいた。

【0195】

本研究は、実施例2および5で評価した同一の17個の臨床的フィーチャ(表4)および43個の分子的フィーチャ(表6)についての情報に基づいた。496個の形態計測的フィーチャセット(表1)を、表18(本明細書中に添付)に示す33個のフィーチャに減少させた。実施例5由来の同一の患者262人についての臨床情報、分子情報、および形態情報を、本研究で評価した。本研究は、上記の実施例5で参照した同一のSVRCフィーチャ減少法を使用した。

【0196】

結果

8個のフィーチャ(実施例5で選択された同一の6個のフィーチャおよび1個のさらなる臨床的・病理学的フィーチャおよび1個のさらなる形態計測的フィーチャ)に基づき、一致指数(CI)0.86の最終モデルを、研究結果として生成した。モデルに含まれる8個のフィーチャを、最終モデルに対するその各フィーチャの寄与と共に図15に示す。本研究で選択したさらなる臨床的フィーチャは、実施例2と併せて上記の生検グリーンソンスコア(-10.60)であった。本研究で選択したさらなる形態計測的フィーチャは、赤色チャンネル(-11.26)で示した間質内のテクスチャの変動であった。このフィーチャは、その選択特性に基づいた間質テクスチャの変動を示し、良性エレメントと対照的に腫瘍に関連した間質の生化学的特質を反映する可能性が最も高い。

【0197】

検証研究

患者366人からなる外部コホートを使用して最終モデルを検証した。最初の5年以内に前立腺癌再発を経験するリスクのある患者の同定を検証するために、CI 0.82、

10

20

30

40

50

感度 96%、および特異性 72% の最終モデルを生成した。以下の表 12 は、訓練および検証コホートについて認められた臨床的フィーチャを示す。表 13 a ~ c は、訓練コホート由来の認められたバイオマーカー - 分子的フィーチャを示す。

【 0 1 9 8 】

【 表 1 2 - 1 】

表 12

特徴	訓練	検証
N	262	61
年齢(年)		
平均	62	61
中央値	63	62
範囲	38-81	42-74
人種		
コーカサス	235 (89.7%)	58 (95.1%)
アフリカ系アメリカ人 (ヒスパニックおよび非ヒスパニック)	21 (8.0%)	2 (3.3%)
他/未知	6 (2.3%)	1 (1.6%)

10

【 0 1 9 9 】

【表 1 2 - 2】

術前PSA (ng/mL)		
平均	10.7	12.9
中央値	7.8	10.0
範囲	0.9-81.9	2.0-69.5
病理学的TNM期		
T2N0	158 (60.3%)	回収せず
T3aN0	70 (26.7%)	回収せず
T3bN0	17 (6.5%)	回収せず
T1-3N+	17 (6.5%)	回収せず
欠測	0	回収せず
UICC期		
T1a < 5%	1 (0.4%)	回収せず
T1b ≥ 5%	1 (0.4%)	回収せず
T1c (明確でも視覚可能でもない)	113 (43.1%)	回収せず
T2a ≤ ½葉	54 (20.7%)	回収せず
T2b ≤ 1葉	43 (16.4%)	回収せず
T2c 兩葉	33 (12.6%)	回収せず
T3a 片側ECE	15 (5.7%)	回収せず
T3c SV+	2 (0.8%)	回収せず
欠測	0	回収せず
直腸の指診の結果		
不明確	122 (46.6%)	32 (52.5%)
明確	140 (53.4%)	29 (47.5%)
欠測	0	0
リンパ節の関与		
陰性	246 (93.9%)	56 (91.8%)
陽性	16 (6.1%)	5 (8.2%)
欠測	0	0
精囊の関与		
なし	233 (88.9%)	51 (83.6%)
あり	29 (11.1%)	10 (16.4%)
欠測	0	0
切除縁		
陰性	216 (82.4%)	36 (59.0%)
陽性	46 (17.6%)	25 (41.0%)
嚢外の関与		
なし	159 (60.7%)	43 (70.5%)
あり	103 (39.3%)	18 (29.5%)
欠測	0	0
優位生検グリーソン分類		
1	1 (0.4%)	0 (0.0%)
2	43 (16.4%)	0 (0.0%)
3	181 (69.1%)	39 (63.9%)
4	36 (13.7%)	22 (36.1%)
5	1 (0.4%)	0 (0.0%)
欠測	0	0

10

20

30

【 0 2 0 0 】

【表 1 2 - 3】

生検グリーソンスコア		
2	1 (0.4%)	0 (0.0%)
3	0 (0.0%)	0 (0.0%)
4	7 (2.7%)	0 (0.0%)
5	56 (21.4%)	3 (4.9%)
6	97 (37.0%)	27 (44.3%)
7	85 (32.4%)	20 (32.8%)
8	13 (5.0%)	8 (13.1%)
9	3 (1.2%)	3 (4.9%)
欠測	0	0
有意検体グリーソン分類		
2	20 (7.6%)	0 (0.0%)
3	198 (75.6%)	34 (55.7%)
4	44 (16.8%)	23 (37.7%)
5	0 (0.0%)	4 (6.6%)
検体グリーソンスコア		
5	21 (8.0%)	1 (1.6%)
6	86 (32.8%)	8 (13.1%)
7	144 (55.0%)	37 (60.7%)
8	11 (4.2%)	7 (11.5%)
9	0 (0.0%)	8 (13.1%)
倍数性		
二倍体	141 (53.8%)	回収せず
四倍体	113 (43.1%)	回収せず
異数性	8 (3.1%)	回収せず
欠測	0 (0.0%)	回収せず
S期の倍数性率(%)		
平均	2.4	回収せず
中央値	1.2	回収せず
範囲	0.0 - 66.4	回収せず
倍数性画分の比率		
平均	3.4	回収せず
中央値	2.4	回収せず
範囲	0.0-20.0	回収せず

10

20

【 0 2 0 1 】

30

【表 1 3 - 1】

表 13a. 組織学的成分による細胞染色比率および染色強度(訓練組)

マーカー	腫瘍			PIN			萎縮性腺		
	1+	2+	3+	1+	2+	3+	1+	2+	3+
Ki-67									
平均±SD	22.0±30.4	7.2±17.1	1.8±4.0	23.0±31.5	7.8±18.3	2.0±4.5	1.3±8.05	1.2±9.9	0.3±1.6
中央値	0.7	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
範囲	0.0-100.0	0.0-100.0	0.0-26.3	0.0-100.0	0.0-100.0	0.0-39.5	0.0-96.0	0.0-96.5	0.0-13.0
CK18									
平均±SD	NA	NA	100.0±0.04	NA	NA	100.0±0.04	NA	NA	100.0±0.00
中央値	NA	NA	100.0	NA	NA	100.0	NA	NA	100.0
範囲	NA	NA	50.0-100.0	NA	NA	50.0-100.0	NA	NA	100.0-100.0
CD45									
平均±SD	NA	NA	0.0±0.04	NA	NA	0.0±0.01	NA	NA	0.0±0.00
中央値	NA	NA	0.0	NA	NA	0.0	NA	NA	0.0
範囲	NA	NA	0.0-0.4	NA	NA	0.0-0.1	NA	NA	0.0-0.0
CD68									
平均±SD	NA	NA	0.0±0.01	NA	NA	NA	NA	NA	NA
中央値	NA	NA	0.0	NA	NA	NA	NA	NA	NA
範囲	NA	NA	0.0-0.1	NA	NA	NA	NA	NA	NA

【 0 2 0 2】

表 13b. 組織学的成分によるCD34染色された細胞の比率(訓練)

	PIN		腫瘍		腫瘍/PIN		PIN/間質	
	間質	腫瘍	腫瘍/PIN	間質	腫瘍/間質	PIN/間質	間質	
平均±SD	0.0±0.04	0.0±0.11	0.1±0.18	0.0±0.07	0.0±0.08	0.0±0.04	0.0±0.04	
中央値	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
範囲	0.0-0.4	0.0-1.7	0.0-0.9	0.0-0.5	0.0-0.4	0.0-0.3	0.0-0.3	

10

20

30

40

【表 1 3 - 2】

表13c. 組織学的成分による染色指数(訓練組)

マーカー	腫瘍	PIN	腺
AR			
平均±SD	171.8 ± 75.9	79.9 ± 83.3	29.5 ± 67.9
中央値	200	66.0	0
範囲	0-300	0-300	0-300
CK14			
平均±SD	2.2 ± 6.4	35.2 ± 62.0	8.3 ± 32.6
中央値	0	0	0
範囲	0-69	0-300	0-300
サイクリンD1			
平均±SD	1.4 ± 7.1	0.0 ± 0.21	0.0 ± 0.0
中央値	0	0	0
範囲	0-90	0-3	0-0
PSA			
平均±SD	117.9 ± 71.2	140.5 ± 97.4	22.4 ± 54.9
中央値	100	134	0
範囲	0-300	0-300	0-300
PSMA			
平均±SD	0.3 ± 2.1	5.8 ± 19.8	3.0 ± 23.2
中央値	0	0	0
範囲	0-21	0-154	0-300
p27^{Kip1}			
平均±SD	3.9 ± 8.2	6.6 ± 19.0	1.3 ± 8.7
中央値	0	0	0
範囲	0-48	0-140	0-120
Her-2/neu			
平均±SD	3.5 ± 16.9	0.2 ± 1.1	0.0 ± 0.0
中央値	0	0	0
範囲	0-150	0-10	0-0

(実施例7) 根治的前立腺摘除後の臨床的な失敗の予測

臨床データ、形態計測データ、および分子データ

本研究は、根治的前立腺摘除を受けた前立腺癌患者の臨床的な失敗を予測するためのモデルを生成した。本研究では、臨床的な失敗を、転移性疾患および/またはアンドロゲン非依存性疾患の発症(例えば、術後のADTの際に転移またはPSAが陽性の骨スキャンが上昇する)と定義した。これは、臨床的な失敗を任意の原因による死亡と定義した実施例2に記載の臨床的な失敗の研究と対照的である。

【0203】

患者345人についての臨床情報および形態情報を、本研究で評価した。表4および6に示す臨床的フィーチャおよび形態計測的フィーチャに類似の10(10)個の臨床的フィーチャおよび27個の形態計測的フィーチャを評価した。11(11)個の分子的フィーチャも評価した。しかし、これらの分子的フィーチャを、コンピュータによって免疫蛍光(IF)検出に供した組織の画像から生成し、実施例2に記載の分子的フィーチャと同様にIHCに基づかなかった。

【0204】

より詳細には、アンドロゲン受容体(AR)、ラセマーゼ(AMACR)、サイトケラチン18(CK18)、TP73L(p63)、および高分子量ケラチンのAlexa蛍光色素標識抗体を、DAPIと共に「quint-plex」アッセイで使用した。蛍光色素の特有のスペクトルプロファイルに基づいて、抗原特異的グレースケール画像を取得した。Definiens Cellenger製品のためのスクリプトを生成して、各抗原を局在化した。抗原の分布および強度から、スクリプトは、細胞型および細胞区画(例えば、管腔上皮細胞、上皮/間質核)を同定し、前立腺腫瘍、良性の腺、および間質中のARおよびAMACRを定量した。すなわち、DAPIおよびCK18画像をそれぞれ有効な核および細胞質のオブジェクトにセグメント化したDefiniens Cellenger製品のスクリプトが生成された。これらのスクリプトは、強度閾値に基づいて画像オブジェクトを核または細胞質に分類した。この強度閾値は、バイオマーカーのグレースケール画像の特徴(例えば、平均強度、標準偏差、変位値)の一次関数であった。各

10

20

30

40

50

スクリプト/一次関数は、所与のバイオマーカーに特異的であり、管理学習（エキスパートによる閾値設定）およびエキスパート閾値と画像特性との間の線形回帰を使用してデザインした。同定した細胞質オブジェクトは、上皮および間質オブジェクトへの核の分割のためのアンカーオブジェクトとしての機能を果たした。ARおよびAMACRバイオマーカーも、空間的および強度の共局在化（colocalization）を使用してセグメント化した。空間的局在化により、以下の特定の区画内のバイオマーカーが同定された：ARについての上皮核および間質核およびAMACRについての上皮細胞（細胞質）。DAPIおよびCK18のために使用した閾値関数に類似の閾値関数により、区画内のバイオマーカーシグナルを有効またはノイズにさらに分類することができた。患者のコアがアレイから欠測された場合、またはサンプルが間質のみを含んでいた場合にその患者を排除した。

10

【0205】

大部分の患者がIFフィーチャを抽出した複数のコアを有していたので、複数のコアを通した各患者およびフィーチャについてのフィーチャ量の集計のための手順が考案された。4つの候補関数（min、max、median、およびmean）を検討した（例えば、患者が3つのコアを有する場合、min関数は3コア間の最も低い値を返す）。所与のフィーチャについて、これらの各関数を適用して、訓練組中の各患者のコア値を集計し、次いで、各aggregation関数について的一致指数を計算して、臨床的な失敗の予測因子として評価した。フィーチャについての最良のaggregating関数を、その一致指数がランダム（0.5）から最も遠いものであると見なした。以下の表14は11個のIFフィーチャおよび対応する選択されたaggregating関数を列挙し、表15はフィーチャを記載する。次いで、選択されたaggregating関数を使用して、各患者の各IFフィーチャについての1つの値を生成した。複数の免疫蛍光に関するさらなる詳細は、2006年10月13日に出願され、発明の名称が「Multiplex In Situ Immunohistochemical Analysis」である共同所有された米国特許出願第11/581,043号（米国特許出願公開第2007/0154958号）（その全体が本明細書中で参考として援用される）に記載されている。

20

【0206】

【表14】

30

表14

IFフィーチャ	選択されたaggregating関数
EpitarposamaMeanMeanChannel30017	mean
EpitarposamaMeanMeanChannel30021	median
Mamacrpostotalarea0045	max
Mcytoplasmamacrpostotalarea0050	max
SumMen02_10totalarea	適用されず;これは、その値が10個の 主なフィーチャの集計値の平均である 誘導フィーチャである
Mepithnucleiarposamacrpostot0085	max
RelAepithNucARpos	median
RelAepithNucARposAMACRpos	max
MeanEpithNucARPosIntensity	min
MeanEpithNucARPosAMACRPosIntens	max
NormSimpleAverageTotalEpithLum	min

40

【0207】

【表 15 - 1】

表15

フィーチャ	説明
EpitarposamaMeanMeanChannel30017	ARバイオマーカーを使用して陽性に発現し、AMACRバイオマーカーに関して陰性の細胞質中に存在する上皮核に分類されるDAPIオブジェクトの平均強度
EpitarposamaMeanMeanChannel30021	ARバイオマーカーを使用して陰性に発現し、AMACRバイオマーカーに関して陽性の細胞質中に存在する上皮核に分類されるDAPIオブジェクトの平均強度
Mamacrpostotalarea0045	上記ノイズ/シグナル閾値を超える強度のAMACRオブジェクトの全領域

10

【 0 2 0 8 】

【表 1 5 - 2】

フィーチャ	説明	
Mcytoplasmamacrposttotalarea0050	AMACRを使用して陽性に発現した細胞質(CK18)の全領域	
SumMen02_10totalarea	<p>上皮核として分類され、ARバイオマーカーを使用して陽性に発現した全DAPIオブジェクトでは、正規化したARオブジェクト強度値 \bar{y}_k の和を以下のように計算する:</p> $I = \sum_k \bar{y}_k$ <p>(式中、$\bar{y}_k = y_k / Y_k$ は、実際のARオブジェクト強度 y_k および局所(AR画像に関して)ノイズ対シグナル強度閾値 Y_k の比である)。I値は、バイオマーカーによって検出されたタンパク質の総量に関連する。I値は、ピニングのための対象である。上皮核を、M=11ピンに分類する。あらゆるピンは、上皮核中の一定量のARを特徴づける。フィーチャは、$20 \leq I < 30$ の上皮核領域である。領域値を、画像上で検出された上皮核の全領域で割る。</p>	10
Mepithnucleiarposamacrpostot0085	ARを使用して陽性に発現し、AMACRに関して陽性である上皮核の全領域。	20
RelAEpithNucARpos	総上皮核領域に関してAR+である上皮核の相対領域。	
RelAEpithNucARposAMACRpos	AR+上皮核に関してAMACR+であるAR+上皮核の画分(相対領域)。	
MeanEpithNucARPosIntensity	Iピン値にわたって平均として計算した上皮核中のARオブジェクトについての合計強度(I値)の値。フィーチャを、画像(目的の領域)中の上皮核中で発現したARの平均量と解釈することができる。	30
MeanEpithNucARPosAMACRPosIntens	強度の平均値によって評価したAR+およびAMACR+上皮核中のARオブジェクトについての合計強度の値(I値)。フィーチャは、AMACRマーカーに関して陽性の細胞質中に存在する上皮核中に発現されたARの平均量と比例する。	

【 0 2 0 9 】

40

【表 1 5 - 3】

フィーチャ	説明	
NormSimpleAverageTotalEpithLum	照度の概念—オブジェクトによって放出された総光エネルギーを導入した。あらゆるDAPIについて、上皮核照度として分類したオブジェクトを計算した(オブジェクト領域上で平均AR強度を掛ける)。フィーチャは、照度値によって計算した上皮核中のAR量の推量である。	

「AMACR陰性である上皮細胞内に存在する核AR」を除く全IFフィーチャは、— 50

致指数 (C I 0 . 4 または C I 0 . 6) に基づいた単変量分析における臨床的な失敗との関連を示した。

【 0 2 1 0 】

結果

本研究に由来するモデルは、7個のフィーチャ(3個の臨床的フィーチャ、1個の分子的フィーチャ、および3個の形態計測的フィーチャ)に基づき、訓練コホートに関して一致指数 (C I) 0 . 9 1、感度 9 5 %、および特異性 8 0 % を有していた。モデル中に含まれる7個のフィーチャを、図 1 6 に示す。前立腺摘除後の臨床的失敗が予測される3個の臨床的フィーチャは、生検グリーンスコア、リンパ節の関与、および検体(前立腺摘除)グリーンスコアであった。3個の形態計測的フィーチャは、青色チャンネル中に示された上皮細胞質の平均強度 (C y t o p l a s m M e a n M e a n C h a n n e l 1 6 0 0 6 0)、赤色チャンネル中に示された間質内のテクスチャの変動 (S t r o m a M e a n S t d d e v C h a n n e l 4 0 3 1 0)、および赤色チャンネル中に示された上皮核間のテクスチャの変動 (E p i t h e N u c l e i S t d d e v M e a n C h a n n 4 0 1 5 7) であった。上記のように、間質テクスチャに関連するフィーチャは、その染色特性に基づき、良性エレメントと対照的に、腫瘍に関連した間質の生物学的特質を反映する可能性が最も高い。さらに、典型的には、細胞質の色が薄青色から濃青色に変化する場合、これは、癌の発症(すなわち、間質領域への上皮細胞浸潤)に起因する組織の変化を反映する。上皮核内側のテクスチャの変動に関して(折り畳み/非折り畳みクロマチンテクスチャおよび核小体)、癌の発症は、典型的には、非折り畳みクロマチンテクスチャを有する上皮核数の増加および核小体数の増加によって特徴づけられ、癌の発症によってこのフィーチャの値がより高くなる。選択された1個の分子的フィーチャは、A M A C R 陽性であった上皮細胞内の A R 強度であった。

10

20

【 0 2 1 1 】

検証

患者 3 1 9 人からなる独立したコホートを使用して最終モデルを検証した。検証のために、前立腺摘除後の臨床的な失敗を予測するための C I 0 . 8 5、感度 8 9 %、および特異性 7 7 % の最終モデルを生成した。以下の表 1 6 は、訓練および検証コホートについての認められた臨床的フィーチャを示す。

【 0 2 1 2 】

30

【表 16 - 1】

表16. 臨床情報

特徴	訓練	検証	
人種			
白人(ヒスパニックおよび非ヒスパニック)	328 (95.1)	301 (94.4)	
アフリカ系アメリカ人(ヒスパニックおよび非ヒスパニック)	11 (3.2)	9 (2.8)	
他/未知	6 (1.7)	9 (2.8)	
術前PSA(ng/ml)			
平均	11.4	11.1	
中央値	7.9	8.2	
範囲	0.5-100.0	1.1-56.2	
臨床的TNM期			
T1a/b	4 (1.2)	4 (1.3)	10
T1c	172 (49.9)	148 (46.4)	
T2a	64 (18.6)	59 (18.5)	
T2b	29 (8.4)	26 (8.2)	
T2c	67 (19.4)	77 (24.1)	
T3	9 (2.6)	5 (1.6)	
リンパ節の関与			
陰性	332 (96.2)	308 (96.6)	
陽性	13 (3.8)	11 (3.4)	
精囊の関与			
なし	316 (91.6)	291 (91.2)	
あり	29 (8.4)	28 (8.8)	
切除縁			
陰性	228 (66.1)	199 (62.4)	20
陽性	117 (33.9)	120 (37.6)	
嚢外の関与			
なし	242 (70.1)	222 (69.6)	
あり	103 (29.9)	97 (30.4)	
優位生検グリーソン分類			
1	2 (0.6)	1 (0.3)	
2	22 (6.4)	20 (6.3)	
3	264 (76.5)	247 (77.4)	
4	56 (19.4)	51 (16.0)	
5	9 (2.6)	0 (0.0)	
生検グリーソンスコア			
2	1 (0.3)	0 (0.0)	
3	2 (0.6)	1 (0.3)	
4	7 (2.0)	12 (3.8)	
5	28 (8.1)	27 (8.5)	
6	186 (53.9)	159 (49.8)	30
7	93 (27.0)	97 (30.4)	
8	23 (6.7)	20 (6.3)	

【 0 2 1 3 】

【表 16 - 2】

9	5 (1.5)	3 (0.9)	
10	0 (0.0)	0 (0.0)	
優位検体グリーソン分類			
2	9 (2.6)	8 (2.5)	
3	269 (78.0)	243 (76.2)	
4	62 (18.0)	63 (19.8)	40
5	5 (1.5)	5 (1.6)	
検体グリーソンスコア			
5	16 (4.6)	16 (5.0)	
6	112 (32.5)	102 (32.0)	
7	182 (52.8)	167 (52.4)	
8	22 (6.4)	18 (5.6)	
9	12 (3.5)	16 (5.0)	
10	1 (0.3)	0 (0.0)	

別の態様では、前立腺摘除後にADTで処置した患者のサブセットの評価に基づいて、漸増レベルのARがADT処置後の臨床的な失敗までの期間の短縮に関連し得ると判断された。したがって、前立腺摘除検体中のAR含有量の測定を使用して、療法のアンドロゲン抑制型に対する応答を予測することができる。まとめると、前立腺摘除検体時および任

意の治療形態の前のAR含有量を使用して、疾患の進行だけでなく、治療に対する潜在的な応答も予測することができる。

【0214】

(実施例8) 前立腺癌再発の予測

臨床データ、形態計測データ、および分子データ

別の研究を行って、根治的前立腺摘除を受けた患者における前立腺癌再発までの期間を予測するモデルを生成した。実施例2、5、および6で、再発までの期間を、根治的前立腺摘除からPSA(生化学的)再発までの期間(月)と定義した。

【0215】

682人の拡大したコホートを評価した。本研究は、実施例6で使用した臨床的フィーチャに匹敵する臨床的フィーチャを使用した(表19を参照のこと)。患者101人(14.8%)は、0.2ng/mL超の2つの連続するPSA測定値として定義されるPSA再発を経験した。認められた再発なしの患者の追跡の中央値は74ヶ月であった。再発までの全時間の中央値に届かなかった。表18に示した同一の形態計測的フィーチャを評価したが、本研究では、Adobe ImageReady 7.0を使用して、デジタル処理で輪郭を描き、その後の画像セグメント化のための各コア内の浸潤性腫瘍のみの領域をカラーマスキングした。定量的多重(quantitative multiplex)免疫蛍光(IF)によって分子的フィーチャを生成した。これは、実施例7と併せて上に一般的に記載されている。IF定量に関するさらなる詳細を、以下に示す。

10

【0216】

モデルのPSA再発を高精度でさらに予測し、一致指数は0.77、感度72%、および特異性74%であった。図17に示すように、AR+上皮核の相対領域に対応する1つの分子的フィーチャを選択した。5個の臨床的フィーチャ(精嚢浸潤、生検グリーンスコア、嚢外拡大、術前PSA、および優位前立腺摘除グリーンソン分類が含まれる)、ならびに腫瘍のみの領域中の2つの形態計測的フィーチャ、上皮核のテクスチャ、および細胞質のテクスチャも選択した。図17では、負の加重(weight)は各フィーチャの存在(または連続的フィーチャのより高い値)がより短いPSA再発までの時間に関連することを示す一方で、正の加重はその逆を示す。これらの加重は、完全なモデル中の各フィーチャの各寄与を示す。

20

【0217】

モデル中に包含するために選択された形態計測的フィーチャは、色のバリエーション(強度の標準偏差)を使用した上皮核領域および細胞質領域のテクスチャ特性を記載する。非折り畳みクロマチンテクスチャおよび上皮核中の核小体外観などの癌進行に関連する生物学的過程により、核画像オブジェクトのテクスチャが変化する。後者により、色標準偏差の値が増加する。細胞質領域のテクスチャ特性は、組織の構造変化および組織が良性期から悪性期に進行した場合の間質への上皮細胞の浸潤に起因する。

30

【0218】

定量的免疫蛍光(IF)

本研究における免疫蛍光画像分析によって定量されたフィーチャを、表14および15に示し、説明している(実施例7を参照のこと)。注目すべきは、ARの強度などのフィーチャが抗原濃度に関連し、(0-255)の8ビット(RGB-(赤色、緑色、青色)値の制限)とは対照的に12ビットを使用して表示されたことである。したがって、ダイナミックレンジの拡大が可能である。ARの測定のうち、1つのフィーチャ(全上皮核の全DAPI領域に対する上皮細胞核(腫瘍および非腫瘍エレメントが含まれる)中のAR量)は、腫瘍中のARについてのIHC染色指数と高度に調和していた(スピアマンの順位相関係数=0.44; p=0.0011)。このフィーチャはまた、一変量で分析した場合にPSA再発と独立して関連した。したがって、定量的免疫蛍光データは、AR IHCバイオマーカーの結果を支持し、細胞区画中のAR発現への値のより客観的な割り当て手段を提供する。

40

【0219】

50

高解像度グレースケール *tiff* 画像 (1280 × 1024 ピクセル) を、TMA コア中の上皮核および間質核内の AR についての強度レベルを同定したオブジェクト指向の方法を使用して個別に処理した。腫瘍エレメントは、正常なエレメントと区別されなかった。むしろ、間質細胞および皮細胞のみが区別された。CK18 発現は、全前立腺上皮細胞のための組織分布マーカーとしての機能を果たした。各「形態学的」層を覆い (すなわち、核については DAPI、上皮細胞および AR については CK18、および、個別の Alexa 蛍光色素、ならびに、バックグラウンドおよび非特異的結合を除去するための予め確立した閾値の使用による)、AR 強度を正確に定義した。多重蛍光画像化によるタンパク質マルチプレックスの自動化局在化および定量のための閾値の確立に関するさらなる説明を以下に示す。この分析により、絶対陽性ピクセル数 (*absolute positive pixel count*) に対するタンパク質濃度が、一連の形態学的相関を反映する各フィーチャ (AR で占められた核 (DAPI 染色) 領域ならびに上皮核および間質核内に存在する AR タンパク質の相対量が含まれる) と同等と見なされた。AMACR 陽性上皮細胞を、AR 含有量についてのさらなる記述子として使用し、誘導されたフィーチャを、全モデル開発の一部として評価した。

10

20

30

40

50

【0220】

図 18 a ~ e は、前立腺組織中の核 AR を定量するための多重画像セグメント化を示す。図 18 a は、個別の形態学的 (すなわち、DAPI、CK18) および AR マップを示す 3 つのグレースケール画像 (図 18 b ~ d) を生成するために分解された標準的な多蛍光色素画像である。図 18 e は、スクリプトによって定量された上皮細胞内の核 AR を示すセグメント化画像である。これらの全ての図を、再現性のために白黒で示す。より詳細には、図 18 a は、核について DAPI で染色した前立腺組織 (示さず)、上皮細胞について Alexa 488 で標識した CK18 (典型的には、緑色で示す)、ならびに間質細胞および上皮細胞中で Alexa 568 で標識した AR (典型的には、赤色で示す) の多重免疫蛍光画像である。各蛍光色素を分離するための *spectral optics* の適用により、核 (図 18 b)、上皮細胞 (図 18 c)、および AR (図 18 d) についての各グレースケール画像を作製した。ピクセルおよびオブジェクト分類に基づいた方法を使用して、カラー合成画像をレンダリングした (図 18 e; ここでは白黒で示す)。これは、CK18 陽性上皮細胞内に存在する DAPI 染色核内に存在する AR および CK18 陰性である核 (間質) 内に存在する AR をセグメント化する。セグメント化した画像は、上皮核中の AR および間質内の AR の定量の基礎であった。計算されたパラメータには、AR に関連する全核領域、平均強度および最大強度、ならびに分布 (陽性細胞の比率) が含まれていた。

【0221】

フィーチャの選択

実施例 8 の拡大した研究では、SVRC Bootstrap フィーチャ選択法を使用した。最初のフィルタリング工程で、目的の転帰と一変量的に相関しなかったフィーチャを除去した。次に、訓練データから N 個の異なるスプリット (本研究では、N = 25) を作成した。各スプリットでは、全訓練インスタンスの約 2/3 を訓練サブセットに無作為に割り当て、全訓練インスタンスの約 1/3 を試験サブセットに無作為に割り当てた。

【0222】

最初のフィルタを通過した全フィーチャから開始する「グリーディ・フォワード」フィーチャ選択過程から方法を開始した。各フィーチャについて、全スプリットにわたる訓練サブセットに対して N 個のモデルを構築し、N 個の各試験サブセットに対して検証した。N 回の実行にわたる各フィーチャについての総合的パフォーマンスを平均した。最良の総合的パフォーマンスを有するフィーチャを選択した。次の工程では、各フィーチャを事前に選択したフィーチャに付加し、再度 N 個のモデルを構築し、スプリットにわたって試験した。その付加によって最良の総合的パフォーマンスが得られたフィーチャを選択した。パフォーマンスが改善されるフィーチャがそれ以上存在しなくなるまで、本方法をこの様式で続けた。

【0223】

その後、「グリーディ・バックワード」フィーチャ選択アプローチを使用した。各フィーチャを除去し、スプリットにわたるそのフィーチャを含まないN個のモデルを構築し、試験した。その除去によって最良の総合的パフォーマンスが得られたフィーチャを除去し、フィーチャの除去によってモデルパフォーマンスが停止して改善されるまで手順を繰り返した。この工程によってモデルの複雑さが軽減され、最初に有意であったかもしれないフィーチャを除去したが、その情報寄与をその後に付加したフィーチャ内にカプセル化した。

【0224】

最後に、完全なSVRCモデルを、完全な訓練データに対して全ての選択したフィーチャを使用して訓練した。

10

【0225】

他の実施形態では、診断的針生検時または根治的前立腺摘除後のいずれかでのアンドロゲン除去療法に前立腺癌が骨転移および/またはPSAに進行する確率を予想するモデルを提供することができる。

【0226】

(実施例9) 根治的前立腺摘除後の臨床的な失敗の予測

臨床データ、形態計測データ、および分子データ

別の研究を行って、根治的前立腺摘除を受けた前立腺癌患者における臨床的な失敗(CF)を予測するためのモデルを生成した。実施例7と同様に、臨床的な失敗を、転移性疾患および/またはアンドロゲン非依存性疾患の発症と定義した。

20

【0227】

32個の臨床的な失敗事象を有する患者373人の訓練組を評価して、管理された多変量機械学習(bootstrappingフィーチャ選択を使用したSVRC)によってモデルを生成した。前立腺摘除後5年以内の臨床的な失敗を予測するために、モデルは、CI 0.92、感度90%、および特異性91%を有していた。モデルを、29個の臨床的な失敗事象を有する患者385人の独立したコホートに対して検証し、CI 0.84、感度84%、および特異性85%を得た。高レベルのアンドロゲン受容体により、アンドロゲン除去療法後の去勢PSAの上昇までの期間がより短くなることが予測された。

【0228】

図19に示すように、モデルは、優位前立腺摘除グリーソン分類およびリンパ節の浸潤状態の臨床的フィーチャならびに管腔周囲の平均長、管腔の相対領域、および緑色チャネル中の細胞質内のテクスチャの変動の形態計測的フィーチャを含む。モデルはまた、定量的多重IFによって生成されたAR+/AMACR-上皮核中のARチャネルの平均強度の分子的フィーチャを含む。図19では、負の加重は各フィーチャの存在(または連続的フィーチャのより高い値)がより短い臨床的な失敗までの時間に関連することを示す一方で、正の加重はその逆を示す。これらの加重は、完全なモデル中の各変数の各寄与を示す。

30

【0229】

本研究では、前立腺摘除前または生物学的再発前であるが前立腺摘除直後のいずれかで治療を受けた患者を排除し、全コホート中に患者881人が残った。コホートを無作為化し、臨床的な失敗事象と類似の数値を有する訓練および検証組に均等に分けた。臨床的な失敗を、転移(骨疾患またはリンパ節または実質臓器における軟組織疾患が含まれる)、去勢状態でのPSAの上昇、または前立腺癌に帰する死亡の明確なX線または病理学的な証拠として事前に特定した。CFまでの期間を、根治的前立腺摘除から第1のこれらの事象までの期間と定義した。患者が最後の来院以後にCFを経験しなかった場合、または最後の来院時の患者の転帰が未知の場合、患者の転帰を打ち切った。

40

【0230】

完全な臨床データ、形態計測データ、および分子データ、ならびに非欠測転帰情報を有する患者のみをさらに研究した。結果として、評価可能患者数は、訓練組で373人、検

50

証組で385人であった(表20および21を参照のこと)。これらの患者758人の特徴は、患者881人の全コホートの特徴と類似していた。実施例8(例えば、定量のために腫瘍エレメントのみを含むように標識化ツール(Adobe 7.0)を使用したコアのデジタル処理によるマスキング)に関連して一般に記載のように、形態計測的フィーチャを患者コアから撮影した画像から抽出した。同様に、実施例8に関連して一般に記載のように、多重定量的IFを使用して分子的フィーチャを生成した。

【0231】

患者373人の訓練組では、33人(8%)が前立腺摘除後にCFを有し、24人が可能な補助療法前に骨スキャンが陽性であり、9人の去勢PSAが上昇した。1人のCF患者を除く全てがホルモン療法を受けていた。ホルモン療法を受けなかった1人の患者は、救済放射線療法を受けていた。さらに、放射線療法およびホルモン療法の両方を受けた患者が1人存在していた。これらの患者373人を、前立腺摘除から76ヶ月後の中央値について追跡し、CFまでの全時間の中央値に届かなかった。モデルによる選択のための変数として9個の臨床的フィーチャを含めた(表20)。72ヶ月の追跡中央値を有する患者385人からなるコホートの検証では、29人(7.5%)の患者がCFを有し、22人が可能な補助療法前に骨スキャンが陽性であり、7人の去勢PSAが上昇した。全CF患者がホルモン療法を受けており、7人が救済放射線療法も受けていた。ブラインドした転帰データと共に6個のフィーチャモデルを適用した。

10

【0232】

各患者のH&E染色腫瘍コアの腫瘍マスキングしたデジタル化画像から、27個の形態計測的フィーチャを生成し(表22)、これは染色の色およびテクスチャならびに前立腺組織サンプル中の上皮細胞および間質細胞の全領域を反映していた。単変量分析では、12個のフィーチャは、各フィーチャの一致指数(CI 0.4またはCI 0.6)に基づいたCFとの関連を示した(表20)。

20

【0233】

特異的細胞型(上皮腫瘍細胞)中のARおよびAMACRを定量して、同時多重抗原評価および画像分析を組み合わせたスペクトル画像分析の使用によって分子的フィーチャを生成した。ARおよびAMACRレベルに関連する免疫蛍光フィーチャのうち、11個の免疫蛍光フィーチャは、単変量分析におけるCFとの関連を示し(CI 0.4またはCI 0.6)、これらを選択のためにモデル中に含めた(表20)。

30

【0234】

最終モデルに関して、前立腺癌進行の予測における腫瘍分化のために同定された役割(すなわち、グリーソン分類)および転移の証拠(すなわち、リンパ節の関与)に注目する。モデルの3個の形態計測的フィーチャは、グリーソン分類に関連する異なる細胞型および全組織構造内の生化学的性質および代謝的性質の代替物としての組織化学的染色の重要性を示す。病理学者は、良性細胞型と反応性または悪性細胞型とを区別する場合の色の重要性を伝統的に承認しているが、色を標準化し、客観的に定量することができなかった。本モデルを使用して、上皮細胞質のテクスチャ色特性は前立腺摘除グリーソン分類よりもCFに強く関連し、疾患進行における細胞の生化学的性質の重要性が示唆される。3個の画像化フィーチャについての各カプラン・マイヤー曲線は、患者を正確に階層化する能力を示した。分子的フィーチャに関して、AMACR陰性上皮細胞内のARの正規化した輝度/強度のみを、モデルに含めるために選択した。このフィーチャについての負の加重(図19)は、AR量の増加がCFまでのより短い時間に関連したことを示す。注目すべきは、術前PSAおよびGleasonの和の情報モデルに含めるために選択されなかったことである。SVRc分析にもかかわらず、これらの変数を、前立腺癌の生化学的性質に関連する分化および色の特徴を反映する画像分析フィーチャと置換した。

40

【0235】

CFモデルの高精度の予測によってより情報量が多くて適切な治療計画が開発可能であると考えられる。これには、再発リスクが高い場合のより早い、ADTの投与(進行性前立腺癌に対して現在において最初に行われる全身療法である)、放射線療法、および/ま

50

たは化学療法の可能性が含まれるであろう。さらに、高リスクの患者の早期同定により、進行を調査する機会が増大する。モデル予測はまた、生化学的再発後の治療決定の誘導で有用であり得る。生化学的再発と臨床的に有意な事象との間の間隔が典型的には数年であり、非常に変動もするので、生化学的に再発した患者によっては迅速な介入が必要ないかもしれない。さらに、生化学的再発と診断された男性は、今日、一連の無痛性疾患を有する可能性が高い。浸潤性疾患（例えば、術後5年以内にCFに進行する疾患）の確率の知識は、患者および患者の主治医による、生化学的再発が積極的療法を誘発するかどうかの決定に役立ち得る。

【0236】

この臨床的進行モデルにおけるARの重要性を考慮すると、前立腺摘除検体中のAR含有量を使用してアンドロゲン除去療法(ADT)後の去勢PSAの上昇までの期間を予測することができるかどうかを決定するための研究を行った。患者881人のうちの63人は事前にCFがなくADTを受けた。これらの患者のうちの32人(51%)がその後CFを示した。患者63人の転帰データを使用して、CFに関連する任意の免疫蛍光フィーチャもADTに対する応答に関するかどうかを調査した。以下の4個のフィーチャは、去勢PSAの上昇までの期間の減少に有意に($p < 0.05$ カイ2乗検定)関連し、その全てが腫瘍上皮細胞中のAR核レベルの増加に関連する。

【0237】

【化1】

<u>フィーチャ</u>	<u>P値</u>
AMACR(-)上皮細胞におけるARの強度	0.0021
AMACR(+)上皮細胞におけるARの強度	0.0003
AMACR(+)上皮細胞における平均AR強度	0.0285
全上皮細胞における総AR強度(正規化)	0.018

腫瘍サンプルのAR含有量の使用により、泌尿器科医はADTの有用性を予測することができ、これは各患者でさえも可能であり、したがって、治療レジメンを調整することができる。現在のCFモデルで高スコアを有し、高レベルのARを有する患者は、以下の特定のジレンマに直面する：ADTおよびおそらく救済放射線療法に対する永続的応答の見込みの減少に伴う臨床的進行の確率の上昇。これらの患者は、ARまたはその補因子をターゲティングするヒストンデアセチラーゼインヒビターなどの療法の良好な候補であり得る。したがって、アンドロゲン受容体の正確な定量が治療決定の誘導を補助すると考えられる。かかるアプローチは、臨床試験の選択をさらに改善し、特にARまたはAR補因子をターゲティングするヒストンデアセチラーゼインヒビターなどの療法のための治療効果の基準としての役割を潜在的に果たすことができる。

【0238】

マルチスペクトル蛍光画像化によるタンパク質マルチプレックスの自動化局在化および定量

免疫蛍光(IF)顕微鏡画像中のタンパク質バイオマーカーの発現の自動化局在化および定量のためのシステムおよび方法を開発した。システムおよび方法は、バイオマーカーシグナルとバックグラウンドとを区別し、ここで、シグナルはIFで使用した任意の多数のバイオマーカーまたは対比染色の発現であり得る。システムおよび方法は、管理学習に基づき、画像バックグラウンド特性の関数としてバイオマーカー強度閾値を示す。前立腺摘除を受けた患者における前立腺癌再発の予測におけるシステムおよび方法の有用性を証明する。具体的には、アンドロゲン受容体(AR)発現を示すフィーチャは、単変量分析における転帰不良と統計的に有意に関連することを示す。ARフィーチャはまた、多変量再発予測に有益であることを示す。

【0239】

緒言

タンパク質バイオマーカーは、癌診断、予後診断、および治療応答の予測のために組織病理学で広く使用されている。これらは、細胞中のタンパク質の発現レベルに関する情報を提供し、それにより、病状に関連する特定の細胞活性が検出可能である。

【0240】

1つのタンパク質発現の定量方法はIFであり、これは、蛍光色素で標識した抗体を標的タンパク質に結合する組織に導入することによってタンパク質を局在化する(21)。特定の波長を有する光源を使用した蛍光顕微鏡下で染色されたスライドを照射する。この励起光は蛍光色素によって吸収されて、より長い波長の光を発する。放射光の強度は、標的タンパク質濃度の基準である。多重化では、組織をいくつかの抗体で同時に標識する。抗体を、異なるスペクトル特性を有する蛍光色素で標識する。複数のバイオマーカーの分離を、組織のマルチスペクトル画像分析およびその後の各抗体の発現を示す画像を得るためのスペクトル分解(spectral unmixing)によって行う。

10

【0241】

伝統的には、組織中の目的のオブジェクト(例えば、核)のその認識された強度レベルに基づいて、IF画像を病理学者によって解釈する。この実務は骨が折れ、観察者内および観察者間の変動に悩まされる。自動化システムの開発により、IF画像の視覚的スコアリングの低コストで客観的な代替法が可能である。

【0242】

バイオマーカーの定量を2段階で行う。第1に、バイオマーカーに関連する区画を検出する。次いで、シグナルを区画内のバックグラウンドから分離する。これらのタスクをしばしば強度閾値化によって行う。対話型閾値(22)~(24)の後、バイオマーカー発現を定量するためのコンピュータ測定をしばしば行う。Gordon et al.(23)は、バックグラウンド由来の3つの標準偏差が閾値を超える連続ピクセル組として単細胞画像中のタンパク質を同定した。画像ヒストグラム様式を対話型閾値化と共にAQUAシステム中のバックグラウンドピクセルの同定のために使用する(24)。Rao et al.(25)は、定量のために目的の区画をマーキングするためにマニュアル上皮核描写を使用する。低シグナル/バックグラウンド比の画像における対話型の方法の制限された再現性およびヒストグラム閾値化の不十分な性能は、既存技術の主な欠点である。

20

【0243】

上記制限に取り組むバイオマーカーの局在化および定量のためのシステムおよび方法が提案される。システムおよび方法は、完全に自動化され、それにより再現可能である。システムおよび方法は、その関連する細胞内区画内のバイオマーカーシグナルを定量し、それにより、システムおよび方法が低シグナル/バックグラウンド比に対してロバストになる。

30

【0244】

システムおよび方法は、バイオマーカーシグナルのバックグラウンドとの識別のためのアプローチに基づく。管理学習を使用して、画像バックグラウンド特性の関数としてのシグナル識別のための強度閾値をモデリングする。前立腺癌の予後診断における提案されたシステムおよび方法の有用性を証明するが、システムおよび方法が多数のさらなる用途を有し得ると理解されるであろう。

40

【0245】

多重画像分析

提案されたシステムおよび方法を、上記のIF多重アッセイと合わせて記載する。IF多重アッセイは、サイトケラチン18(CK18)、 α -メチルアシル-CoA-ラセマゼ(AMACR)、およびアンドロゲン受容体(AR)と共に核対比染色4'-6-ジアミジノ-2-フェニルインドール(DAPI)を含む。

【0246】

画像分析プラットフォーム

Definiens Enterprise Image Intelligence Suite(商標)を使用して定量化システムをデザインした(26)。有益な生物学

50

的オブジェクト（例えば、核）への画像分割は、オブジェクト指向の画像分割に基づく多段階過程である。このパラダイムでは、オブジェクトは、ピクセルよりもむしろ画像処理操作およびフィーチャ計算が好ましい最小単位である。例えば、強度閾値を画像に適用する場合、オブジェクト強度を閾値に供する。オブジェクトの強度は、オブジェクトに属する全ピクセルの平均強度である。

【0247】

2つの方法を使用して、プリミティブを得る（26）。多重解像度分割法により、ピクセルとオブジェクトの形状規則性との間の色類似性に基づいてプリミティブが見出される。四部木分割法は、色類似性のみを使用する。スケールパラメータは、両方法においてオブジェクトのサイズを調節する。

10

【0248】

バイオマーカーの局在化および定量

発展したアプローチでは、第1の工程は、対応する細胞内区画内のバイオマーカーの空間的局在化である。例えば、上皮核および細胞質は、それぞれ、ARおよびAMACRの細胞内区画である。次の工程は、対応する区画プリミティブ中の真のバイオマーカーシグナルとバックグラウンドとを識別することである。バックグラウンドは、自己蛍光および蛍光色素の組織への非特異的結合からなる。細胞内区画のオブジェクトは、オブジェクトが一定量の真のシグナルを含む場合、バイオマーカーに関して陽性に分類される。バイオマーカーの定量化は、分類したオブジェクトの性質（例えば、領域）を測定する分析の最終工程である。

20

【0249】

バイオマーカー閾値モデル

画像分析の次の段階は、強度閾値を使用した真のバイオマーカーシグナルとバックグラウンドとの識別である。閾値 T を、バイオマーカーのバックグラウンド特性の関数と見なす。この関数を見出すために、オブジェクト B を以下の2つの非重複クラスに分割する：区画 B_c および区画の外側のオブジェクト B_b 中の局在化。オブジェクト B_b は、以下のフィーチャによって説明されるバックグラウンドと見なす：平均 μ_b 、標準偏差 σ_b 、およびオブジェクト強度の t_h パーセンタイル q_b 、（式中、 $\mu_b = 5, 10, \dots, 95$ ）。閾値 T とバックグラウンド特性との間の線形関係を推測する場合、

$$T = \beta^T X \quad (1)$$

（式中、 $\beta = [\beta_0, \beta_1, \dots, \beta_p]^T$ はモデルパラメータであり、 $X = [1, X_1, \dots, X_p]^T$ はバックグラウンドフィーチャである）。(1)では、 T は各画像の性質にしたがって変化することに留意のこと。

30

【0250】

多変量線形回帰を使用して、 T を決定する。モデル訓練のために画像組を選択する。各画像について、a) オブジェクト B_c および B_b を同定し、b) バックグラウンドフィーチャを抽出し、c) 病理学者がエキスパート強度閾値 $expert_T$ を決定する。訓練組を使用してモデルを訓練し、エキスパート閾値($expert_threshold$)を予測する。

【0251】

モデル過剰適合を回避するために、冗長且つ統計的に有意でないフィーチャを訓練から排除した。2個のフィーチャのうち1個は、対応のある相関関数 r が $r < 0.85$ を満たす場合に冗長である。エキスパート閾値とのより強い一変量相関を有するフィーチャを、多変量回帰のために使用する。回帰係数のための信頼区間を計算し、その信頼区間がゼロを含む係数を統計的に有意でないとして見なしして消去する。パラメータ β を、残存するフィーチャを使用して算定する。

40

【0252】

局在化のための細胞内区画

基本細胞内組織区画は、標的ARおよびAMACR抗原を起源とする真のシグナルを含む。上皮核および細胞質は、上記基本区画を構成する。これらは、それぞれDAPI画像

50

およびCK18画像からセグメント化された陽性オブジェクトである。次の工程は、細胞質との重複によって核オブジェクトを上皮細胞および間質細胞に分割することである。

【0253】

癌再発予測では、浸潤性癌領域中に存在する上皮核を識別することが重要である。後者は、AMACR発現を含む細胞質領域として認識される。AMACR - と対照的に、これらの上皮核をAMACR + と呼ぶ。

【0254】

最初のプリミティブオブジェクトを、CK18およびAMACR画像のカラー四部木分割によって得る。これらのオブジェクトを、多重解像度分割を使用してより粗いオブジェクトに再セグメント化する。次いで、表23由来の閾値(3)を適用して、最初の細胞質またはCK18+オブジェクトを同定する。全ての最初のCK18オブジェクトにわたって(3)中の強度標準偏差bを算定することに留意のこと。一旦最初のCK18+オブジェクトが同定されると、近傍分析を使用して、クラスラベルをリファインする。その全境界長に対する細胞質の境界長が0.6以上である小さなバックグラウンドオブジェクトを、細胞質に再分類する。小さな単離細胞質オブジェクトを、バックグラウンドとして標識する。

10

【0255】

プリミティブオブジェクトを、DAPI画像でオーバーレイする。表23由来の閾値(2)を使用して、DAPI+オブジェクトおよびバックグラウンドオブジェクト中のプリミティブを分類する。この分類を、上記に類似の近傍分析によってさらにリファインする。DAPI+オブジェクトは、DAPI画像中に示された実際の核の小フラグメントである。

20

【0256】

領域拡張およびオブジェクト分類からなる対話型方法を使用して、DAPI+プリミティブを核にマージする。最後に、CK18+オブジェクトとの共同在化に基づいて、核オブジェクトを、上皮核および間質核に分類する。

【0257】

陽性および陰性バイオマーカー発現

AMACR+細胞質領域は、以下の2つの工程で見出される：a) AMACR画像上のプリミティブオブジェクトのCK18+領域でのオーバーレイ；b) 表23中の閾値(5)の共同在化オブジェクトへの適用。その強度が(5)を超える細胞質オブジェクトを、AMACR+に分類する。それぞれ陽性および陰性細胞質とのその関連に応じて、上皮核を、AMACR+およびAMACR-と標識する。

30

【0258】

上皮細胞は、AR発現の基本的細胞内区画である。表23中の閾値モデル(4)を、上皮核中に共同在化したARオブジェクトに適用する。そのAR+オブジェクト領域がその全領域の0.1%を超える場合、上皮核をAR陽性(AR+)に分類する。そうでなければ、AR陰性(AR-)に分類する。

【0259】

バイオマーカー定量化

一旦バイオマーカーが共同在化および分類されると、オブジェクトベースの強度および領域特性を示す定量的フィーチャを、共同在化バイオマーカーから生成する。代表的な強度フィーチャには、画像中で認識されたAMACR陽性/陰性上皮細胞内の平均AR強度および総AR強度が含まれる。代表的な領域フィーチャには、全上皮細胞領域と比較したAR+上皮細胞領域および全AR+細胞間のAMACR+上皮細胞画分が含まれる。

40

【0260】

結果

IF画像分析のために提案されたシステムの有用性を、前立腺摘除を受けた患者の前立腺癌再発の予測で証明した。予測のために2つの有害転帰型を考慮した。一方の転帰は、患者の血清PSAレベルの有意な上昇に対応する前立腺特異的抗原(PSA)再発(PS

50

A R)であった。他方の転帰は、遠隔転移などの基準によって測定した有意な疾患の進行に対応する臨床的な失敗 (C F)であった。

【 0 2 6 1 】

それぞれ P S A R および C F 予測タスクにおける患者 6 8 2 人および 7 5 8 人について、組織領域の少なくとも 8 0 % が腫瘍で覆われた組織マイクロアレイコアを得た。コアを D A P I 対比染色で標識し、C K 1 8、A R、および A M A C R バイオマーカーを C R I N u a n c e (商標) マルチスペクトル画像分析システムを使用して画像化した (2 7)。各コアについて、D A P I 対比染色の発光ピーク波長での 1 つの 1 2 ビット 1 2 8 0 × 1 0 2 4 ピクセルのグレースケール画像を取得した。C K 1 8、A R、および A M A C R バイオマーカーのそれぞれについて、対応する蛍光色素の発光スペクトルを含む一定範囲の波長についての画像スタックを取得した。得られた画像スタックを N u a n c e (商標) システムを使用して分解し、それぞれバイオマーカーの 1 つに対応する 3 つの画像を得た。

10

【 0 2 6 2 】

訓練組由来の 6 0 個の多重画像組に関して、D A P I 対比染色のエキスパート閾値、ならびに C K 1 8、A R、および A M A C R バイオマーカーを認定された病理学者が選択し、閾値モデルを上記および表 2 3 に示すように訓練した。モデルを、そのエキスパートと予測される閾値との間の最大の矛盾を有する画像の目視検査によって検証した。画像上の A R および A M A C R バイオマーカーを定量し、全上皮核領域と比較した A R + 上皮核領域ならびに A R + 上皮核の強度を示すフィーチャを得た。A R + A M A C R + および A R + A M A C R - 上皮核についての類似のフィーチャも算定した。

20

【 0 2 6 3 】

利用可能な患者を、P S A R 予測のために患者 3 4 2 人および 3 4 0 人の訓練および検証組にグループ分けし、C F 予測のために患者 3 7 3 人および 3 8 5 人の訓練および検証組にグループ分けした。P S A R 予測では、A R + 上皮核の相対領域は、訓練組において最も高い (全 A R フィーチャの間で) 一変量一致指数 (C I) 0 . 3 7 を示した (p 値 < 0 . 0 0 1)。C F 予測では、A R + A M A C R + 上皮核の強度は、最も高い一変量 C I 0 . 3 0 を達成した (p 値 < 0 . 0 0 1)。

【 0 2 6 4 】

さらに、多変量モデルを、システム病理学パラダイムに照らして 2 つの予測タスクのために訓練した。このパラダイムでは、患者の臨床的プロフィール (例えば、年齢)、組織学的プロフィール (ヘマトキシリンおよびエオシン染色組織検体から抽出した画像フィーチャによる (2 8))、および分子学的プロフィール (タンパク質バイオマーカーの発現を測定するフィーチャによる) 由来の異種情報を管理学習フレームワーク中で組み合わせて、癌再発を予測した。得られた P S A R モデルは、上の実施例 8 に記載されている。得られた C F モデルは、上の実施例 9 に記載されている。

30

【 0 2 6 5 】

したがって、I F 多重化顕微鏡画像中の任意のタンパク質バイオマーカー発現の自動化局在化および定量化のためのシステムおよび方法を本明細書中に提供する。管理学習問題としてのバイオマーカー画像の閾値アプローチ、画像バックグラウンド特性の関数としての閾値のモデル化によって真のバイオマーカーシグナルをバックグラウンドと識別する。例えば、前立腺癌予後診断アッセイの I F 多重成分へのこのシステムの適用を証明する。A R 発現の領域および強度を示すフィーチャは、有意な一変量相関を有することが示された。さらに、多変量環境での他の予測因子と競合して有益である新規、再生可能、およびロバストなフィーチャを選択し、それにより、実際の適用におけるシステムの有用性が検証される。

40

【 0 2 6 6 】

【表 2 3 - 1】

表23. 多重バイオマーカーのための閾値モデル

バイオマーカー	モデル	Eq.
DAPI	$T = 0.86\sigma_b + 1.1q_{b,5} - 32.0$	(2)
CK18	$T = 0.55 \sigma_b + 220$	(3)
AR	$T = 5.1q_{b,5} + 0.39q_{b,95} + 1100$	(4)

【 0 2 6 7 】

【表 2 3 - 2】

AMACR	$T = 2.1\mu_b - 0.82q_{b,25} + 1800$	(5)
-------	--------------------------------------	-----

10

(実施例 10) ゲフィチニブで処置した非小細胞肺癌患者の生存の予測；臨床データ、形態計測データ、および分子データ

別の研究を行って、ゲフィチニブで処置した非小細胞肺癌 (NSCLC) を有する個体の生存を予測するモデルを生成した。本発明の前に、かかる予測を行うための正確な分析ツールは存在しなかった。上記のように、本発明のシステム病理学アプローチは、PSA再発および前立腺摘除後の臨床的な失敗を正確に予測することが示された。本研究は、本発明を使用して、ゲフィチニブ処置後の生存を正確に予測することもできることを証明する。

20

【 0 2 6 8 】

種々の固形腫瘍 (非小細胞肺癌 (NSCLC) が含まれる) 中の上皮成長因子受容体 (EGFR) の発現の豊富さが、EGFRを選択的分子治療薬 (詳細には、ゲフィチニブ (IRESSA) などのチロシンキナーゼインヒビター) の魅力的な標的にしている。予め処置した患者におけるゲフィチニブについての最初の結果が有望であるにもかかわらず、2つの他の薬剤と組み合わせたゲフィチニブの2つの極めて重要な第III相試験では、ゲフィチニブで処置したNSCLC患者は、プラセボ群と比較して、有意により良好な全生存率や無増悪生存率を示さなかった。さらに、第III相 (肺癌 (ISEL) 試験におけるプラセボ対照 IRESSA 生存評価) では、ゲフィチニブ単剤療法は、全生存におけるいくつかの改善と関連した。しかし、データは統計的有意性に到達しなかった。人種 (アジア)、性 (女性)、喫煙歴 (非喫煙)、良好なパフォーマンスステータス、および腫瘍病理学 (腺癌 - 気管支肺胞パリアント) などの臨床 - 人工統計学フィーチャを有するので、ゲフィチニブに対する正の反応は、EGFRのATPポケット変異の活性化に関連した。これらの結果により、NSCLC患者の特定の下位集団がゲフィチニブに応答する可能性が最も高いことが示唆された。しかし、データは、低EGFR変異頻度 (具体的には、コーカサス集団；米国に約10~15%) によって複雑になっており、(制限された研究における) ゲフィチニブで処置した患者の約45~50%が幾らかの臨床的利点を経験していたという所見と対照的である。別の課題は、ゲフィチニブ対未処置および種々の化学療法に無作為化した患者の間のEGFR変異に関する比較データの欠如である。さらに、応答率は、研究した群の人口統計 (例えば、性比、民族性、および詳細な喫煙歴) に非常に依存するようである。それにより、応答基準が以前に考慮していたよりも複雑であることが示唆される。重要なことに、プラセボ対照 ISEL 試験由来の最近の証拠により、高EGFR遺伝子コピー数によって臨床的利点および生存が予測されることが示唆される。これらの所見は、応答基準を理解するために認められたいくつかの複雑さを示す。

30

40

【 0 2 6 9 】

患者109人のコホートを評価した。具体的には、ゲフィチニブで処置した進行性難治性NSCLC由来の診断腫瘍サンプルを評価した。腫瘍サンプルを、EGFR DNA変異分析、EGFR免疫組織化学、組織学的形態計測、および15種のマーカーについての定量的免疫蛍光によって評価した。サポートベクトル回帰数学モデルを使用して、6個の臨床的フィーチャ (性、喫煙歴、診断時の年齢、腫瘍病歴、およびECOGパフォーマンス

50

スステータス)を組織学および定量的バイオマーカー多重化と統合した。

【0270】

患者および組織

第1のコホートは、284人のゲフィチニブ250mgで毎日経口にて処置した進行性難治性NSCLCを有する米国人患者からなる。以下の5個の臨床的フィーチャを分析した：性、喫煙歴、診断時の年齢、病歴、ECOGパフォーマンスステータス(スケールは、0(健康)から5(病死)までの範囲である)(表24)。非染色スライド(細針吸引、細胞ペレット、およびサイトスピンが含まれる)および/または診断検体由来のパラフィンブロックを、ヘマトキシリンおよびエオシン(H&E)を使用して腫瘍含有量について評価した。全バイオマーカーを、臨床転帰の知識を用いずに分析した。

10

【0271】

EGFRチロシンキナーゼドメイン変異分析

各パラフィンブロック由来の2つの連続的な厚さ20μmの切片またはパラフィンスライド由来の最小の8個の非染色切片を分析した。プロテイナーゼKとのインキュベーション後に脱パラフィン組織サンプルからゲノムDNAを得て、その後クロロホルム抽出およびエタノール沈殿を行う。EGFR変異を、最初のエクソン19、20、および21のDNA配列決定ならびにエクソン21のL858R点変異および最も一般的なエクソン19欠失(del G2235 - A2249)を検出するための第2の増幅困難な変異系(amplification refractory mutation system)(ARMS)、具体的には、対立遺伝子特異的ポリメラーゼ連鎖反応(PCR)の使用によって分析した。腫瘍DNA由来の少なくとも2つの独立したPCR産物におけるARMSまたは順方向および逆方法の両方での遺伝子配列決定のいずれかによってチロシンキナーゼドメインの変異が検出された場合、患者を変異陽性に見なした。

20

【0272】

組織学的形態計測

H&E染色スライドを、元のブロックまたは非染色切片から調製した。代表的な腫瘍領域由来の1~6つの画像を、20倍のOlympus明視野顕微鏡と共にSPOT Insight QEカラーデジタルカメラ(KAI2000)を使用して取得した。画像分析ソフトウェアにより、画像オブジェクトを赤色、緑色、および青色チャンネル値を示す組織病理学的細胞エレメント、一般的形状フィーチャ(例えば、領域、長さ)、および空間的関係の性質(例えば、全組織に関する管腔の相対量)に分類し、これらから統計を行った。サンプル調製(すなわち、サイトスピン、針生検、および組織切除)、固定、染色、および組織の質の相違により、画像セグメント化のためのいくつかの異なるスクリプトが開発された。

30

【0273】

EGFR免疫組織化学

EGFRタンパク質を、EGFR pharmDXキット(DAKO, Glostrup, Denmark)を使用した免疫組織化学によって分析した。10%超の腫瘍細胞が染色された場合にサンプルを陽性に見なした。各強度レベル(0~3)をその強度レベルでの細胞の比率と乗じることによって各サンプルの染色指数を計算した。したがって、この指数は、0から300までの範囲である。

40

【0274】

(多重(M-plex(商標))バイオマーカー評価)

サイトケラチン18、Ki67、カスパーゼ-3活性化、CD34、リン酸化EGFR、リン酸化ERK、リン酸化AKT、PTEN、サイクリンD1、リン酸化mTOR、PI3-K、VEGF、VEGFR-2、およびリン酸化VEGFR-2が含まれる15種の抗体を選択し、6つの多重形式にグループ化した。各抗体により、形態学的成分(例えば、サイトケラチン18(CK18)、CD34)、シグナル伝達経路、または増殖、アポトーシス、および血管形成の生体過程を評価した。

【0275】

50

各抗体を、免疫組織化学 (I H C) によって一連の細胞株および / またはコントロール肺癌組織サンプルについて評価した。各抗体を、免疫組織化学および免疫蛍光の標準的な操作手順の両方を使用して滴定した。次いで、単一および多重の免疫蛍光形式に抗体を進めた。15種の抗体を、生物学的適用に関して抗原回復系、アイソタイプ、および優先順位づけに基づいて6個の多重形式にまとめた。これらの m - p l e x 形式を、表 27 に列挙する。

【 0 2 7 6 】

【 表 2 7 】

表27. 抗体および多重免疫蛍光分類

M-plex群	抗体	製造供給元	カタログ番号	アイソタイプ	希釈	蛍光標識
1	サイトケラチン18	Vector	VC-C414	マウスIgG1	1:250	1° - 488
	サイクリンD1	BioCare Med.	CP236B	ウサギIgG	1:100	1° - 568
2	サイトケラチン18	Vector	VC-C414	マウスIgG1	1:250	1° - 488
	Ki-67	Ventana	790-2910	マウスIgG1	1:75の前希釈	2° - 555
	PI3キナーゼ	Cell Signaling	3821	ウサギIgG	1:20	1° - 568
	活性化カスパーゼ-3	Chemicon	ab3623	ウサギIgG	1:5	2° - 594
3	サイトケラチン18	Vector	VC-C414	マウスIgG1	1:200	1° - 488
	CD-34	Dako	M7165	マウスIgG1	1:50	2° - 555
	p-mTOR	Cell Signaling	ab2971	ウサギIgG	1:50	1° - 568
	pERK	Cell Signaling	ab4376	ウサギIgG	1:5	2° - 594
4	サイトケラチン18	Vector	VC-C414	マウスIgG1	1:200	1° - 488
	pAKT S473	Abcam	ab4802	ウサギIgG	1:20	1° - 555
	PTEN	NeoMarkers	MS1797	マウスIgG1	1:50	2° - 568
	KDR	Upstate	07-158	ウサギIgG	1:450	2° - 594
5	サイトケラチン18	Vector	VC-C414	マウスIgG1	1:250	1° - 488
	VEGF	Abcam	ab1316	マウスIgG1	1:50	2° - 555
	pKDR	Upstate	07-374	ウサギIgG	1:150	2° - 594
6	サイトケラチン18	Vector	VC-C414	マウスIgG1	1:10	2° - 488
	EGFR	Dako	K1492	マウスIgG1	前希釈	1° - 555
	pEGFR (Y1068)	Abcam	ab5644	ウサギIgG	1:250	1° - 594

組織サンプルの脱パラフィンおよび再水和を、標準的な操作手順によって行った。電子レンジ中の 1 x R e v e a l 溶液 (B i o C a r e M e d i c a l) 中でのスライドの 7 . 5 分間のボイルによって抗原回復を行った。スライドを室温で 2 0 分間冷却し、次いで、リン酸緩衝化生理食塩水 (P B S) 中にて 3 分間の洗浄を 2 回行った。

【 0 2 7 7 】

組織サンプルに対して以下のプレハイブリッド形成処置工程を行った。組織の細胞構造の浸透を補助するために、サンプルを、P B T (P B S + 0 . 2 % T r i t o n X - 1 0 0) 中にて室温で 3 0 分間インキュベートし、その後 P B S 中にてそれぞれ 3 分間で 3 回のリンスを行った。組織自己蛍光の減少を補助するために、サンプルを、酸アルコール (1 % H C l を含む 7 0 % エタノール) 中にて室温で 2 0 分間インキュベートし、その後 P B S 中にてそれぞれ 3 分間で 3 回のリンスを行った。スライドの I F ブロッキング試薬 (0 . 5 m g / m l B S A を含む P B S) 中での室温で 2 0 分間のインキュベーションによって非特異的結合部位をブロッキングした。ブロッキング工程とその後のハイブリッド形成工程との間に洗浄を行わなかった。

【 0 2 7 8 】

予め決定した力価を使用して、1つのマウス I g G 1 抗体および 1 つのウサギ I g G 抗体のカクテルを I F ブロッキング試薬中で作製した。約 1 0 0 μ l のこの抗体カクテルを組織サンプルに適用し、抗体および組織サンプルを、湿室中にて室温で 1 時間ハイブリッド形成させた。ハイブリッド形成後、それぞれ P B S 中で 5 分間のリンスを 2 回行い、その後それぞれ P B S 中で 3 分間のリンスを 2 回行った。

【0279】

標識工程のために、上記ハイブリッド形成抗体のための適切な Zenon Alexa Fluor ウサギ IgG 標識および適切な Zenon Alexa Fluor マウス IgG1 標識 (Invitrogen / Molecular Probes) のカクテルを、製造者が推奨する濃度 (各 Fab フラグメントについて 50 倍希釈) にて IF ブロッキング試薬中で作製した。約 100 μ l のこの標識カクテルを組織サンプルに適用し、組織サンプルを湿室中にて室温で 30 分間インキュベートした。標識反応後、2 回リンスした。上記のハイブリッド形成および標識工程を、マルチプレックス中に含まれる次の各抗体群にて繰り返した。2 回までのさらなる抗体ハイブリッド形成および標識ラウンドが可能である。異なる抗体を識別するために、ウサギまたはマウスのいずれかに特異的な Zenon Alexa Fluor 標識を 1 回を超えて使用しなかったことに留意すべきである。

10

【0280】

10%ホルマリン中での室温で 10 分間のサンプルのインキュベーションおよびその後の PBS 中で 3 分間の各リンスを 2 回行うことによって固定工程を行った。DAPI マウンティング溶液を含む約 25.0 μ l の Slow Fade Gold antifade 試薬をサンプルに適用し、その後にカバーガラスをかけた。サンプルを、分析を行うことができるまで、-20 で保存した。

【0281】

Nikon 90i 自動化蛍光顕微鏡をマウントし、MetaMorph オンラインソフトウェアによって制御される CRI Nuance マルチスペクトルカメラ (Cambridge Research and Instrumentation, Inc.) を使用して多重蛍光画像を取得した。得られた画像を定量的グレースケール TIFF 画像 (1280 x 1024 ピクセル) として保存した。目的の選択領域について、DAPI 核対比染色を、帯域通過 DAPI フィルタ (Chroma) を使用して、波長 480 nm で記録した。Alexa 488 を、FITC フィルタ (Chroma) を使用して、10 nm 間隔にて 520 nm と 560 nm との間で取得した。Alexa 555、568、および 594 を特注のロングパスフィルタ (Chroma) を使用して 10 nm 間隔にて 570 nm と 670 nm との間で記録する一方で、Alexa 647 を第 2 の特注のロングパスフィルタ (Chroma) を使用して 10 nm 間隔にて 640 nm と 720 nm との間

20

30

【0282】

画像分析スクリプトを形態学的特質 (すなわち、核については DAPI、上皮細胞については CK18) にオーバーレイし、Alexa 蛍光色素サインを使用して各抗体を区別しながら、バックグラウンドを除去して信号対ノイズ比を最適にする。強度プロフィールによってフィーチャ (各標識抗体についての平均、最大、および標準偏差が含まれる) を生成する。その活性化形態を有する各マーカーの特異的関係 (例えば、pKDR : KDR、pEGFR : EGFR) および別のマーカーを発現する細胞中の特異的マーカーの強度 (例えば、pERK 陽性上皮細胞におけるリン酸化 mTOR) を、スクリプトフィーチャ

40

【0283】

抗体の品質管理

pEGFR (Y1068) 抗体および pERK 抗体の特異性を確認するために、A431 NSCLC 細胞株を上皮成長因子 (EGF) で処置した。細胞を氷冷 PBS で 2 回洗浄し、次いで、フェニルメチルスルホニルフルオリド (1 mmol/L)、ロイペプチン (25 μ g/mL)、アプロチニン (25 μ g/mL)、およびペプスタチン A (3.5 μ g/mL) を含む溶解緩衝液 (0.5% Triton X-100、50 mmol/L Tris (pH 7.6)、300 mmol/L NaCl、2 mmol/L Na₃VO₄) にて 4 で 25 分間溶解した。免疫沈降についての標準的なプロトコール後、I

50

P E C L Plus 化学発光ウェスタンプロットを、E G F 活性化細胞株および非活性化細胞株を使用して行った。V E G F で活性化した H U V E C 細胞株において p K D R を使用して類似の試験を行った。さらに、E G F および V E G F の両方のために、処置細胞および非処置細胞を、日常的な I H C のために処理した。

【0284】

免疫蛍光フィーチャの生成

各抗体の免疫蛍光スクリプト：コントロール肺癌組織サンプルを使用して抗原対を構築した。各タンパク質の予想される細胞分布にしたがってスクリプトをデザインし、取得したグレースケール画像を使用して各マーカーについての領域およびその後の蛍光強度の両方を同定する。予備スクリプトによって400個超のフィーチャ（その多くが冗長）を生成し、それにより、モデル中のエンタリーのためのフィーチャが84個に減少した。

10

【0285】

統計分析

死亡までの期間を、ゲフィチニブでの処置の開始から死亡日または最後の追跡日までの期間と定義し、これを週で報告した。最後の追跡日に死亡が記録されなかった患者を打ち切った。フィーチャ減少した打ち切りデータについてのサポートベクトル回帰 (S V R c - F R) を、M - P l e x (商標)、組織学的形態計測、および E G F R 免疫組織化学によって生成された臨床的変数およびフィーチャに適用した。S V R c は、打ち切りデータ対応のための伝統的 S V R の適応である。これを達成するために、右打ち切りデータおよび非打ち切りデータを処理可能な修正損失 / ペナルティ関数を定義した。S V R c - F R アルゴリズムは、一致指数 (C I) に基づいて目的の転帰と一変量的に相関しないフィーチャを除去するための最初のフィルタリング工程を有する。C I は、2人の無作為に選択した患者の時間（両方が死亡した場合）または打ち切られた患者の最後の追跡前に死亡した患者の時間を、正確な様式で予測する確率である。C I は0から1までの範囲であり、0.5は無作為な関連 (r a n d o m a s s o c i a t i o n) を示す。したがって、C I が0.5から離れるほど転帰とのより強い関連を示す。C I < 0.4 または > 0.6 のフィーチャのみを保持した。C o x モデルを、複数のフィーチャ (S V R c - F R によって選択されるフィーチャが含まれる) のコンパレータとして使用した。段階的アプローチを使用するとき、フィーチャの組み入れ基準は p 値 0.15 であった。モデル中にとどめるために、フィーチャの p 値は 0.05 でなければならなかった。

20

30

【0286】

結果

284人の患者のうち、109人(38%)由来の組織サンプルは、さらなる分析に十分な腫瘍 (> 50%) を有していた。109人の患者サンプルのうち、87人から E G F R 変異の結果を得た一方で51人の患者は最終モデルに含めるための完全なデータ (すなわち、臨床的、免疫組織化学、および免疫蛍光) を有していた。全コホート (患者284人) および最終モデル (患者51人) についての人口統計学を表24に含む。

【0287】

【表 2 4】

表24. 全患者コホートと比較した最終予測モデルにおける患者の特徴

特徴	モデル中の患者 (51人:N(%))	全コホート (患者284人:N(%))
パフォーマンスステータス		
0	2 (3.9)	13 (4.58)
1	39 (76.5)	203 (71.48)
2	10 (19.6)	68 (23.94)
性		
女性	22 (43.1)	126 (44.37)
男性	29 (56.9)	158 (55.63)
病歴		
腺癌	31 (60.8)	138 (48.59)
気管支肺癌	1 (2.0)	5 (1.76)
大細胞癌	5 (9.8)	36 (12.68)
混合癌	2 (3.9)	20 (7.04)
NSCL	5 (9.8)	30 (10.56)
扁平上皮細胞癌	7 (13.7)	55 (19.37)
喫煙歴		
なし	8 (15.7)	45 (15.85)
あり	43 (84.3)	239 (84.15)
化学療法レジメン		
0	1 (2.0)	19 (6.69)
1	19 (37.3)	106 (37.32)
2または3	25 (49.0)	135 (47.54)
4, 5,または6	6 (11.8)	24 (8.45)
年齢		
平均 +/- SD	60.8 +/- 12.83	63.02 +/- 11.29
中央値	62.0	64.0
範囲	24.0 - 82.0	24.0 - 86.0
生存		
死亡	50 (98.0)	268 (94.37)
打ち切り	1 (2.0)	16 (5.63)
生存期間の中央値(週)	25.0	27.0

注釈:連続フィーチャの平均値は類似している。

E G F R 変異状態

エクソン19~21、4、10中の変異についての公開された配列に基づいたE G F R変異分析を、臨床的転帰データおよびDNA分析に十分な材料を有する患者87について行った。患者4人(5%)は、エクソン19、20、または21中にE G F Rチロシンキナーゼドメインの変異を有していた。これらの患者のうちの2人はエクソン19にインフレーム欠損(de1L747-S752insVおよびde1L747P753insS)を有し、1人はエクソン20に体細胞変異(T790M)を有し、1/4はエクソン21にL858R変異を有していた。E G F R変異を有する患者4人全員が女性であり、腺癌と診断された。これらの患者のうちの2人が部分的に応答し、2人が進行した(表25)。短命(4週間)の患者に存在するT790M変異は、薬物耐性と関連すると以前に示されていた。変異について分析した全患者についての全生存率の中央値は24週間であり、1年間の全生存率は26%であった。低頻度のE G F R変異のために、変異状態を、その後のモデル開発における変数として使用しなかった。

【0288】

【表 2 5】

表25. EGFR変異を有する患者についての応答状態

患者数	エクソン	変異	客観的応答	死亡までの期間(週)
46	19	del L747-S752insV	部分的応答	65
127B	19	Del L747-P753insS	部分的応答	92
69	21	L858R	進行	22
196	20	T790M	進行	4

10

EGFR免疫組織化学

腫瘍が50%を超えて存在する患者60人のサンプルに対して免疫組織化学を行った。簡潔に述べれば、腫瘍のみの領域におけるEGFRタンパク質の不均質を評価し、14サンプル(23%)が染色指数200以上を示す一方で(発現の増加が示唆される)、20サンプル(30%)の染色指数が0であった。腫瘍染色指数は、予測モデル中に含まれる唯一の変数であった。EGFRタンパク質の免疫組織化学を陽性細胞率×強度(0~3+)を使用して分析し、分析のために利用可能なサンプルを使用して患者についての染色指数(Hスコア)を作製した(表28)。

20

【0289】

【表 2 8 - 1】

表28. EGFR免疫組織化学由来の詳細な結果

プロトコール番号					
DM05010					
DMスライド番号	総陽性細胞 (%)	+細胞 (%)	++細胞 (%)	+++細胞 (%)	Hスコア*
1	95	25	35	35	200
2	80	30	30	20	150

30

【0290】

【表 28 - 2】

プロトコール番号 DM05010					
DMスライド 番号	総陽性細胞 (%)	+細胞(%)	++細胞(%)	+++細胞(%)	Hスコア*
3	100	0	10	90	290
4	10	10	0	0	10
5	10	5	5	0	15
6	0	0	0	0	0
7	80	10	40	20	150
8	60	20	40	0	100
9	0	0	0	0	0
10	100	0	20	80	280
11	75	35	25	15	130
12	100	0	10	90	290
13	30	20	10	0	40
14	40	10	20	10	80
15	100	10	40	50	240
16	0	0	0	0	0
17	0	0	0	0	0
18	90	10	40	40	210
19	40	15	15	10	75
20	90	20	30	40	200
21	40	10	20	10	80
22	0	0	0	0	0
23	0	0	0	0	0
24	0	0	0	0	0
25	100	0	0	100	300
26	0	0	0	0	0
27	90	20	20	50	210
28	40	20	10	10	70
29	0	0	0	0	0
30	0	0	0	0	0
31	100	0	0	100	300
32	90	10	10	70	240
33	30	10	20	0	50
34	0	0	0	0	0
35	70	20	50	0	120
36	0	0	0	0	0
37	50	20	20	10	90
38	0	0	0	0	0
39	20	20	0	0	20
40	0	0	0	0	0
41	30	10	10	10	60
42	30	10	10	10	60
43	90	20	60	10	170
44	0	0	0	0	0
45	40	30	10	0	50
46	0	0	0	0	0
47	10	10	0	0	10

10

20

30

40

【 0 2 9 1 】

【表 2 8 - 3】

プロトコール番号 DM05010					
DMスライド 番号	総陽性細胞 (%)	+細胞(%)	++細胞(%)	+++細胞(%)	Hスコア*
48	0	0	0	0	0
49	90	10	10	70	240
50	70	30	30	10	120
51	70	10	40	20	150
52	90	10	10	70	240
53	10	10	0	0	10
54	60	50	10	0	70
55	0	0	0	0	0
56	30	10	10	10	60
57	0	0	0	0	0
58	0	0	0	0	0
59	90	10	20	60	230
60	10	0	10	0	20
61	80	10	60	10	160

*Hスコア=(+++とスコアされた細胞の%)×3+(++とスコアされた細胞の%)×2+(+とスコアされた細胞の%)×1。Hスコアの最大は300。

H & E 画像分析

109人のH&E染色患者サンプル由来の選択されたデジタル化画像を、画像ソフトウェアを使用して処理した。本ソフトウェアは、各組織(細胞要素、分離上皮核、上皮細胞質、間質、および肺胞内腔)をセグメント化および分類する。エレメントを、スペクトルの「色」特性、形状、および組織オブジェクト間の空間的關係を使用して分類する。管腔を、例えば、肺胞「腫瘍」上皮細胞に囲まれた余白の全関連から分類する。39個の画像化フィーチャを生成し、予測モデルに入れた。

【0292】

表29は、各患者サンプル(全切片、針生検、および針吸引が含まれる)から抽出した39個のH&Eフィーチャを列挙する。

【0293】

【表 2 9 - 1】

表29. 組織学的形態計測フィーチャ

フィーチャ名称	説明
CytoplasmMeanMeanChannel40058	全細胞質オブジェクトのピクセル値の平均(赤色チャンネル)
CytoplasmMeanMeanChannel50059	全細胞質オブジェクトのピクセル値の平均(緑色チャンネル)
CytoplasmMeanMeanChannel60060	全細胞質オブジェクトのピクセル値の平均(青色チャンネル)
CytoplasmMeanStddevChannel40066	全細胞質オブジェクトのピクセル値の標準偏差の平均(赤色チャンネル)
CytoplasmMeanStddevChannel50067	全細胞質オブジェクトのピクセル値の標準偏差の平均(緑色チャンネル)
CytoplasmMeanStddevChannel60068	全細胞質オブジェクトのピクセル値の標準偏差の平均(青色チャンネル)
CytoplasmStddevMeanChannel40081	全細胞質オブジェクトのピクセル値の平均のSD(赤色チャンネル)
CytoplasmStddevMeanChannel50082	全細胞質オブジェクトのピクセル値の平均のSD(緑色チャンネル)
CytoplasmStddevMeanChannel60083	全細胞質オブジェクトのピクセル値の平均のSD(青色チャンネル)
EpithelialNucleiMeanAreaPx10093	全上皮核オブジェクト領域の平均
EpithelNucleiMeanMeanChanne40104	全上皮核オブジェクトのピクセル値の平均の平均(赤色チャンネル)
EpithelNucleiMeanMeanChanne50105	全上皮核オブジェクトのピクセル値の平均の平均(緑色チャンネル)
EpithelNucleiMeanMeanChanne60106	全上皮核オブジェクトのピクセル値の平均の平均(青色チャンネル)
EpitheNucleiMeanStddevChann40112	全上皮核オブジェクトのピクセル値のSDの平均(赤色チャンネル)

【0294】

10

20

30

40

【表 29 - 2】

EpitheNucleiMeanStddevChann50113	全上皮核オブジェクトのピクセル値のSDの平均(緑色チャンネル)
EpitheNucleiMeanStddevChann60114	全上皮核オブジェクトのピクセル値のSDの平均(青色チャンネル)
EpitheliaNucleiStddevAreaPx10116	全上皮核オブジェクト領域のSD
EpitheNucleiStddevMeanChann40127	全上皮核オブジェクトのピクセル値の平均のSD(赤色チャンネル)
EpitheNucleiStddevMeanChann50128	全上皮核オブジェクトのピクセル値の平均のSD(緑色チャンネル)
EpitheNucleiStddevMeanChann60129	全上皮核オブジェクトのピクセル値の平均のSD(青色チャンネル)
StromaNucleiMeanAreaPx10300	全間質核オブジェクト領域の平均
StromaNucleiMeanMeanChannel40311	全間質核オブジェクトのピクセル値の平均の平均(赤色チャンネル)
StromaNucleiMeanMeanChannel50312	全間質核オブジェクトのピクセル値の平均の平均(緑色チャンネル)
StromaNucleiMeanMeanChannel60313	全間質核オブジェクトのピクセル値の平均の平均(青色チャンネル)
StromaNucleiMeanStddevChann40319	全間質核オブジェクトのピクセル値の標準偏差の平均(赤色チャンネル)
StromaNucleiMeanStddevChann50320	全間質核オブジェクトのピクセル値のSDの平均(緑色チャンネル)
StromaNucleiMeanStddevChann60321	全間質核オブジェクトのピクセル値のSDの平均(青色チャンネル)
StromaNucleiStddevAreaPx10323	全間質核オブジェクト領域のSD
StromaNucleiStddevMeanChann40334	全間質核オブジェクトのピクセル値の平均のSD(赤色チャンネル)
StromaNucleiStddevMeanChann50335	全間質核オブジェクトのピクセル値の平均のSD(緑色チャンネル)
StromaNucleiStddevMeanChann60336	全間質核オブジェクトのピクセル値の平均のSD(青色チャンネル)
AreaCytodivTotTissueArea	全組織領域で除した細胞質領域
AreaEpitNucdivTotTissueArea	全組織領域で除した上皮核領域
AreaLumcndivTotTissueArea	全組織領域で除した管腔領域
AreaLymphdivTotTissueArea	全組織領域で除したリンパ領域
AreaStrNucdivTotTissueArea	全組織領域で除した間質核領域
AreaStromadivTotTissueArea	全組織領域で除した間質領域
NormTotalNucleiArea	全組織領域で除した全核単位領域
NormGlandTissueArea	全組織領域で除した全腺単位領域

SDは標準偏差

10

定量的 M - P l e x (商 標) 免疫蛍光

20

コントロール組織サンプルを使用して評価した場合、示した予想される細胞の局在および分布において15種の全抗体を使用した。患者59人由来の腫瘍検体を、方法の項に列挙した15種の抗体を使用して評価した。各抗体についての目的の代表的領域を、サンプル内の腫瘍の量および質に基づいて選択し、壊死領域および細胞残屑を回避した。各多重分析において患者サンプルあたり最少で3つの視野を取得し、全画像を免疫蛍光ソフトウェアを使用して処理して84個の定量的フィーチャを生成した。活性化カスパーゼ-3およびリン酸化mTORは、浸潤リンパ球および腫瘍上皮細胞内に局所的に見出された。対照的に、サイクリンD1は、腫瘍上皮細胞の核内に支配的に存在した。興味深いことに、pKDRは腫瘍上皮細胞および腫瘍内皮細胞の島内に散在的に見出された。

【0295】

30

84個のフィーチャ(表30)は、見出される区画(すなわち、DAPI=核;サイクリンD1、Ki67、PTEN、およびリン酸化ERK(pERK))については核;CK18、カスパーゼ-3a、およびVEGFについては細胞質;リン酸化AKTについては核/細胞質;EGFRおよびKDRについては膜;ならびにリン酸化KDR(pKDR)、リン酸化mTOR、PI3K、およびリン酸化EGFRについては細胞質/膜)、シグナルの分布(すなわち、核、細胞質などのオブジェクト内の幾何学的ドットとしてのピクセル対ピクセル)、および相互の関係(例えば、EGFR-pEGFR;p-mTOR-pERK)に関する各バイオマーカーからのアルゴリズム出力(強度および領域の両方に関連する平均、最大、および標準偏差)を示す。

【0296】

40

【表 30 - 1】

表30. 定量的免疫蛍光フィーチャ

フィーチャ名称	
Dapiwithcycl0002	
Dcyclkmean500007	
Dcyclvar1000013	
Ocyclmean500027	
Ocyclvar1000030	
Ocyclvar500032	
Oquantile95cycl0038	
Ratio100cyc0041	
Ratio50cyc0043	
B1EgfrEgfrmean500004	
B1EgfrEgfrq950010	10
B1EgfrEgfrvar1000016	
B2pEgfrmean500023	
B2pEgfrvar1000035	
CorrB1EgfrEgfrB2pEgfr0040	
OB1Egfrmean500053	
OB1Egfrq950059	
OB1Egfrvar1000065	
OB2pEgfrmean500072	
OB2pEgfrq950077	
OB2pEgfrvar1000083	
VaroverlapB1Egfr0094	
VaroverlapB2pEgfr0095	
B1Vegfmean500004	
B1Vegfq950010	
B1Vegfvar1000016	
B2pKdrmean500023	20
B2pKdrq950029	
B2pKdrvar1000035	
CorrB1VegfB2pKdr0040	
OB1Vegfmean500053	
OB1Vegfvar1000065	
OB2pKdrmean500072	
OB2pKdrq950077	
OB2pKdrvar1000083	
VaroverlapB2pKdr0095	
B3Caspasemean1000040	
B3Caspasemean500042	
B3Caspaseq950048	
B3Caspaseratio1000049	
B3Caspaseratio500051	
B3Caspasevar1000054	
B4Pi3kmean500061	30
B4Pi3kq950067	
B4Pi3kratio1000068	
B4Pi3kratio500070	
B4Pi3kvar1000073	
Corrb1B2Ki670078	
OB2Ki67mean500122	
OB2Ki67q950127	
OB2Ki67ratio1000128	

【 0 2 9 7 】

【表 3 0 - 2】

OB2Ki67ratio500130
OB2Ki67var1000133
VarB3Caspase0144
VarB4Pi3k0145
B3KDRmean500042
B3KDRq950048
B3KDRratio1000049
B3KDRratio500051
B3KDRvar1000054
B4pAKTmean500061
B4pAKTq950067
B4pAKTratio1000068
B4pAKTratio500070
B4pAKTvar1000073
Corrb1B2pTEN0078
OB2pTENmean500122
OB2pTENq950127
OB2pTENratio1000128
OB2pTENratio500130
OB2pTENvar1000133
VarB3KDR0144
VarB4pAKT0145
MeanVesselArea0030
StdVesselArea0034
EpitInPmtoWithPerkMinMeaCha40026
Dpmtormean500036
Dperkmean500039
Dperkvar1000042
Dpmtormean0048
Dpmtorquantile950056
Dpmtorvar1000058
Dquantile95perk0065
Dvarperk0066

10

20

予測モデルの開発

統合的「システム病理学」アプローチを使用して、全生存の8つのモデルを、臨床的変数とともに以下のフィーチャの異なる組み合わせに基づいて開発した：EGFR免疫組織化学（分子）、組織学的形態計測（形態計測）、および免疫蛍光（分子）。患者によってはいくつかのフィーチャドメインについて欠測データが存在していたので、さらなるドメインを付加した場合、含まれる患者数が最大284人から51人の最少組に減少することに留意のこと。6個の臨床的フィーチャのうちでCIフィルタを通過したものはなかったので、SVRC-FRは、臨床的フィーチャのみを有するデータセットに適用できなかった（患者284人のコホート）。

30

【0298】

【表 3 1 - 1】

表31. SVRC-FRによって構築された8モデル中で選択された性能およびフィーチャ

データセット	患者数	患者生存数	最良モデルのCI	選択したフィーチャ (モデルにおいて重要な順)	フィーチャの寄与
臨床的	284	16	NA	CIフィルタを通過したフィーチャなし	NA
臨床的 + IHC	55	1	0.62	ps_m	-5.35
臨床的 + IF	56	2	0.72	ps_m	-3.85

40

【0299】

【表 3 1 - 2】

				B3Caspasemean1000040	-3.22
				Ocyclvar1000030	-2.29
				B1Vegfvar1000016	-2.24
				VaroverlapB2pKdr0095	-1.99
臨床的 + IHC + IF	52	1	0.73	ps_m	-4.75
				VaroverlapB2pKdr0095	-2.63
				B3Caspasemean1000040	-2.52
				Dcyclkmean500007	-2.21
臨床的 + H&E	53	2	0.67	ps_m	-4.72
				AreaLumendivTotTissueArea	2.46
臨床的 + IHC + H&E	51	1	0.68	ps_m	-4.77
				AreaLumendivTotTissueArea	2.51
臨床的 + IF + H&E	53	2	0.75	ps_m	-3.50
				Dcyclkmean500007	-2.49
				AreaLumendivTotTissueArea	2.36
				B3Caspasemean1000040	-1.89
				VaroverlapB2pKdr0095	-1.86
				EpitInPmtoWithPerkMinMeaCha40026	-1.80
				B2pKdrvar1000035	-1.54
				Ocyclvar1000030	-1.54
				VarB3Caspase0144	-1.41
臨床的 + IHC + IF + H&E	51	1	0.74	ps_m	-4.77
				B3Caspasemean1000040	-2.51
				Ocyclvar1000030	-2.32
				B2pKdrvar1000035	-2.17
				Dcyclkmean500007	-2.08
				AreaLumendivTotTissueArea	1.91

10

20

システム病理学アプローチの有用性を評価するために、上記の 8 つのモデルを、全フィーチャドメインの完全な情報を有する患者のサブセット (n = 51) 由来のデータを使用して訓練した。IHC が CI フィルタを通過しなかったため、このフィーチャの排除によってモデルは変化しなかった (例えば、臨床的 + IHC モデルは臨床モデルに等しい)。表 3 2 は、患者 51 人のコホートに対する SVRc - FR の結果を示す。パフォーマンスステータスは、CI フィルタを通過した唯一の臨床的変数であった。選択された唯一の H & E フィーチャは、管腔の相対領域であった。臨床的フィーチャへの IF フィーチャの付加により、CI が改善された。臨床データへの H & E フィーチャによっても CI が改善されたが、IF よりわずかに低かった。

30

【 0 3 0 0 】

【表 3 2 - 1】

表32. 完全な情報を有する患者51のコホートから構築したSVRc-FRモデル

データセット	最良モデルのCI	選択したフィーチャ (モデルにおいて重要な順)	フィーチャの寄与
臨床的	0.62	ps_m	-4.766
臨床的 + IF	0.73	ps_m	-4.766
		B3Caspasemean1000040	-2.4688
		Ocyclvar1000030	-2.2744

40

【 0 3 0 1 】

【表 3 2 - 2】

		B2pKdrvar1000035	-2.1261
		Dcyclkmean500007	-2.0558
臨床的 + H&E	0.68	ps_m	-4.766
		AreaLumendivTotTissueArea	2.5081
臨床的 + IF + H&E	0.74	ps_m	-4.766
		B3Caspasemean1000040	-2.5079
		Ocyclvar1000030	-2.3232
		B2pKdrvar1000035	-2.1663
		Dcyclkmean500007	-2.0788
		AreaLumendivTotTissueArea	1.9107

10

異なるモデルについてのCoxの結果を、表33に示す。これらは、変数の段階的選択を使用したCoxモデルに基づいた多変数分析である。p値 0.15であった全フィーチャをCoxモデルの入力として入れた。モデル中に留まらせるために、フィーチャはp値 0.05を必要とした。

【0302】

1を超えるハザード比は生存に及ぼす負の影響を示し（フィーチャの値が高いほどより短い寿命が予測される）、1未満のハザード比はフィーチャが保護効果を有することを示す（フィーチャの値が高いほどより長い寿命が予測される）。これらの結果に基づいて、パフォーマンスステータスは、生存の有意に臨床的な予測因子であることが見出された。IFフィーチャのうち、Coxの結果およびSVRCモデル中でしばしば同一の抗体を選択した。しかし、選択したフィーチャは、SVRC-FRを使用して選択した特質と異なる特質を捕捉した。H&Eフィーチャのうち、Coxモデルにより、上皮核の平均領域および細胞質の相対領域で生存が予測されることを見出された。

20

【0303】

【表 3 3 - 1】

表33. Coxの結果

データセット	DF	パラメータ 推定値	標準誤差	カイ2乗	Pr > ChiSq	HR
臨床的						
ps_m	1	0.73059	0.12598	33.6295	<0.0001	2.076
sex_m	1	0.30755	0.12428	6.1234	0.0133	1.360
臨床的 + IHC						
ps_m	1	1.11823	0.32541	11.8089	0.0006	3.059
臨床的 + IF						
ps_m	1	2.76611	0.55914	24.4737	<0.0001	15.897
sex_m	1	-0.90167	0.39445	5.2254	0.0223	0.406
smk_hist_m	1	2.51627	0.55970	20.2116	<0.0001	12.382
Ocyclmean500027	1	0.00399	0.00125	10.1507	0.0014	1.004
B1EgfrEgfrq950010	1	-0.00143	0.0003890	13.5689	0.0002	0.999
B1Vegfmean500004	1	-0.0005077	0.0001250	16.4936	<0.0001	0.999
B1Vegfvar1000016	1	0.00108	0.0004208	6.5637	0.0104	1.001
OB2pKdrmean500072	1	0.0004915	0.0001789	7.5490	0.0060	1.000
B3Caspaseq950048	1	-0.00166	0.0005254	9.9609	0.0016	0.998
Corrb1B2Ki670078	1	-5.34023	1.10152	23.5036	<0.0001	0.005
VarB3Caspase0144	1	0.00643	0.00170	14.2993	0.0002	1.006
Corrb1B2pTEN0078	1	1.56428	0.76844	4.1439	0.0418	4.779
OB2pTENvar1000133	1	-0.01540	0.00423	13.2270	0.0003	0.985

30

40

【0304】

【表 3 3 - 2】

Dperkmean500039	1	0.00140	0.0003034	21.1775	<0.0001	1.001
臨床的 + IHC + IF						
ps_m	1	1.38468	0.41388	11.1932	0.0008	3.994
Dcyclvar1000013	1	0.00164	0.0005454	9.0105	0.0027	1.002
B1Vegfmean500004	1	-0.0001949	0.0000782	6.2195	0.0126	1.000
EpitInPmtoWithPerkMinMeaCha40026	1	0.00378	0.00174	4.7192	0.0298	1.004
臨床的 + H&E						
ps_m	1	1.10707	0.33782	10.7395	0.0010	3.025
smk_hist_m	1	1.70880	0.49665	11.8381	0.0006	5.522
EpithelialNucleiMeanAreaPx10093	1	-0.00958	0.00329	8.4502	0.0037	0.990
臨床的 + IHC + H&E						
ps_m	1	1.25964	0.35019	12.9383	0.0003	3.524
smk_hist_m	1	1.59729	0.49377	10.4646	0.0012	4.940
CytoplasmMeanMeanChannel40058	1	-0.01083	0.00550	3.8834	0.0488	0.989
EpithelialNucleiMeanAreaPx10093	1	-0.01150	0.00349	10.8336	0.0010	0.989
臨床的 + IF + H&E						
Dcyclvar1000013	1	0.00200	0.0005022	15.8982	<0.0001	1.002
OB2pKdrvar1000083	1	0.0006294	0.0002266	7.7146	0.0055	1.001
OB2pTENratio500130	1	-21.08066	8.03902	6.8764	0.0087	0.000
臨床的 + IHC + IF + H&E						
ps_m	1	1.48368	0.41363	12.8661	0.0003	4.409
smk_hist_m	1	1.04319	0.47848	4.7534	0.0292	2.838
Dcyclvar1000013	1	0.00186	0.0006154	9.1377	0.0025	1.002
B1Vegfmean500004	1	-0.0002506	0.0000807	9.6380	0.0019	1.000
EpitInPmtoWithPerkMinMeaCha40026	1	0.00398	0.00184	4.6619	0.0308	1.004
AreaCytodivTotTissueArea	1	2.55288	0.99480	6.5855	0.0103	12.844

DF、自由度;HR、ハザード比

4 個の全フィーチャドメインに基づいたモデルの詳細な説明を、以下に記載する。各入力ドメインについての完全なデータは、患者 51 人に利用可能である。その人口統計学的特徴は、全コホートと類似していた (表 2 4)。

【0305】

最終モデル中で選択されたフィーチャを、モデルにおいて重要な順に表 2 6 に列挙する。フィーチャに対する正の寄与は値が高いほど良好な転帰 (より長い寿命) が予想されることを意味し、負の寄与は値が高いほど悪い転帰が予想されることを意味する。その寄与がゼロから最もかけ離れているフィーチャを、最も重要と見なす。

【0306】

CI 0.74 で最終モデルを実行した。比較すると、臨床データのみを使用したモデルの CI は 0.62 であった。パフォーマンスステータスは、選択された唯一の臨床的フィーチャであった。したがって、免疫蛍光および H & E フィーチャの付加により、生存予測能力は 12% 増加した。

【0307】

モデルスコアを、ログランク検定に基づいたカイ 2 乗統計量によって分析した。この分析により、切点 39.5 は、その実際の生存経験に関して患者を最も強く分離する (カイ 2 乗、21.39; 調整 p 値 0.0002) ことが証明された。39.5 を超えるスコアは、ハザード比 5.26 を有していた (95% 信頼区間: 2.60、10.62)。すなわち、39.5 を超えるスコアを有する患者について、より低いスコアを有する患者と比較して、研究期間にわたる死亡リスクは 5 倍であった。死亡リスクについてのこのハザード比を、遡及的な独立したコホートで確認する必要があるであろう。切点を超える患者および切点未満の患者についての Kaplan-Meier 曲線を、図 2 0 に示すように生成した。

【0308】

各フィーチャの予測値を、Cox 分析によってさらに試験した。この分析により、パフォーマンスステータス (ハザード比: 4.4、95% 信頼区間 1.96 ~ 9.92) が全モデルにわたって生存の有意な予測因子であることが確認された。パフォーマンスステータスよりもあまり確実に実行されない IF フィーチャについて、しばしば同一のマーカーを選択したが (すなわち、サイクリン D 1、pKDR、カスパーゼ - 3)、マーカーの異なる特質を反映する起源の異なるフィーチャを、異なるモデルの間または異なるモデル内で選択した。

10

20

30

40

50

【 0 3 0 9 】

【 表 2 6 】

表26. 臨床的、組織学的形態計測、免疫組織化学的、および免疫蛍光
フィーチャドメインから構築した最終モデル中で選択されたフィーチャ

フィーチャ	寄与
パフォーマンスステータス	-4.766
カスパーゼ-3平均強度レベル	-2.5079
サイクリンD1全強度	-2.3232
pKDR変動性	-2.1663
サイクリンD1平均強度	-2.0788
管腔(肺胞)空間の相対領域	1.9107

10

考察

画像化研究で客観的応答が認められていない場合でさえ、ゲフィチニブから恩恵を受けることができるNSCLC患者の同定が決定的に重要性であることが医学界で同意されている。この理由のために、本発明者らは、終点として全生存を使用してゲフィチニブに対する応答を調査した。ゲフィチニブで処置した患者の全生存のための予測モデルが患者の腫瘍サンプル中の情報の付加によって改善されることが証明された。定量的バイオマーカー特性および組織学的形態計測の臨床データとの統合により、かかる予測モデルの精度(CI)が62%(臨床的フィーチャのみ)から74%(全ドメイン)まで増加した。

【 0 3 1 0 】

20

多数の変数が、ゲフィチニブ感受性に影響を及ぼすと報告されており、変数には、性、腫瘍病歴(例えば、腺癌/気管支肺胞癌)、喫煙歴、民族性、EGFR遺伝子コピー数、EGFRキナーゼドメイン変異、およびEGFR変異型が含まれる。最近、EGFRエクソン19欠失患者がL858R変異を有する患者よりも長い生存期間の中央値を有することが報告された。さらに、EGFRプロモーターにおける生殖系列多形は、チロシンキナーゼドメイン変異を持たない患者におけるゲフィチニブ応答に潜在的に関連する因子として提案されている。さらに、ゲフィチニブに対する応答は、腫瘍病歴または患者の人口統計学的と無関係にNSCLC患者で認められている。

【 0 3 1 1 】

本EAPコホートでは、患者87人のうちの4人にエクソン19~21中にEGFR変異が同定された。変異頻度が低すぎて予測モデルにこれらの変異を含められなかったにも関わらず、変異頻度は主に西洋人の患者を含む以前の研究と一致した。このコホートでは、以前の研究の場合、EGFR変異頻度は、ゲフィチニブ由来の臨床的利点の推定頻度よりもはるかに低かった。本EAPコホート由来の患者124人の分析では、45%の患者が第1の再評価で進行の証拠がなく、29%がゲフィチニブの投与中に肺癌関連症状の改善が報告された。

30

【 0 3 1 2 】

生存に関連すると一貫して同定された唯一の臨床的変数はパフォーマンスステータスであった。これは、良好なパフォーマンスステータス(ECOGステータス0~2)がより長い生存と関連することを証明した以前の研究と一致する。良好な臨床転帰と関連していた気管支肺胞病歴を本発明者らの予測モデルで選択しなかったが、選択した組織学的画像化フィーチャである「総組織領域で除した管腔領域」を、この腺癌サブタイプの代理形態学的フィーチャと解釈することができる。

40

【 0 3 1 3 】

進行性NSCLC患者の下位個体群の異なるゲフィチニブ感受性の基本とする分子機構の理解は不完全なままである。EGFR変異およびEGFR遺伝子コピー数に加えて、試験された他の分子変数には、インタクトなEGFR、リン酸化EGFR、および下流分子(P13K、pAkt、およびpERKなど)が含まれる³¹、⁵²~⁵⁴。不運なことに、幾らかでもある場合にこれらの各因子がゲフィチニブ感受性で果たす役割と一致せず、全生存についてのその重要性とあまり一致しない。免疫組織化学によって評価されたE

50

GFRでさえ、ゲフィチニブ応答とのその関連についての矛盾する報告の主題であった。

【0314】

本分析では、全生存に関連した定量的免疫蛍光によっていくつかのマーカーを同定した。有意には、増加量がより短い全生存と関連した2つのサイクリンD1フィーチャを、最終モデルによって選択した。

【0315】

モデルによって選択された残存バイオマーカー（活性化カスパーゼ-3およびリン酸化KDRが含まれる）を、肺癌臨床検体において広く調査しなかった。カスパーゼ-3は、生存期間の中央値を有する過剰発現に関連する矛盾する報告の主題であった。しかし、pKDRについて、臨床NSCLCサンプルにおける免疫組織化学または免疫蛍光に関する研究は最近発表されていない。Coxモデルにより、（選択されたフィーチャに基づいた）VEGFの過剰発現が生存の減少と関連することが確認され、これは、微小管密度および腫瘍浸潤と併せてVEGFを試験する以前の研究と一致する。サイクリンD1、カスパーゼ-3、pKDR、およびVEGFを種々のサブモデルで繰り返し選択したという事実により、腫瘍の成長、転移、および生存のための細胞周期の調節および血管形成の重要性が示唆される。

10

【0316】

したがって、ゲフィチニブで処置した進行性NSCLC患者において全生存が首尾よく予測された一連の統合的なシステムベースのモデルを構築した。選択したバイオマーカーは、これらが細胞周期の調節、アポトーシス、および血管形成の兆候を示すという点で誘発的である。

20

【0317】

さらなる実施形態

したがって、医学的状態の発生の予測のための方法およびシステムを提供することが分かる。特定の実施形態を本明細書中で詳細に開示しているが、これは、説明のみを目的として一例として開示しており、以下の添付の特許請求の範囲を制限することを意図しない。特に、本発明者らは、特許請求の範囲によって定義された本発明の精神および範囲を逸脱することなく種々の置換、変更、および修正を行うことができることを意図する。他の態様、利点、および修正形態が以下の特許請求の範囲の範囲内であると見なされる。示した特許請求の範囲は、本明細書中に開示の本発明を代表する。他の特許請求の範囲に記載されていない発明も意図される。出願人は、以下の特許請求の範囲中にかかる発明を追求する権利を留保する。

30

【0318】

コンピュータシステムの少なくとも一部を使用して上記の本発明の実施形態を実行可能な範囲で、記載の方法および/または記載のシステムの少なくとも一部の実行のためのコンピュータプログラムが本発明の態様の一部として予測されることが認識されるであろう。コンピュータシステムは、任意の適切な装置、システム、またはデバイスであり得る。例えば、コンピュータシステムは、プログラミング可能なデータ処理装置、汎用コンピュータ、デジタルシグナルプロセッサ、またはマイクロプロセッサであり得る。コンピュータプログラムをソースコードとして具体化してコンピュータ上での実行のためにコンパイルすることができるか、例えば、オブジェクトコードとして具体化することができる。

40

【0319】

前述のコンピュータプログラムまたはコンピュータシステムに帰するいくつかまたは全ての機能性を、例えば、1つまたは複数の特定用途向け集積回路によってハードウェア中で実行することができることも考えら得る。

【0320】

適切には、コンピュータプログラムを、コンピュータ使用可能形態のキャリア媒体に保存することができる。この媒体も本発明の1つの態様として想定される。例えば、キャリア媒体は、固体メモリ、光メモリ、または光磁気記憶装置（読み出し可能なおよび/または書き込み可能なディスク（例えば、コンパクトディスク（CD）またはデジタル多用途

50

ディスク（DVD）など）、または磁気メモリ（ディスクまたはテープなど）であり得る。コンピュータシステムは、システムを操作するためのプログラムを使用することもできる。コンピュータプログラムを、電気信号（高周波搬送波または光搬送波が含まれる）などのキャリア媒体中で具体化される遠隔光源から供給することもできる。

【0321】

（参考文献）

以下の参考文献（このうちのいくつかは上で引用されたものである）は全て、その全体が本明細書中に参考として援用される。

【0322】

【化2】

10

[1] Scherr D., *et al.*, Urology. 61 (2 Suppl 1): 14-24, Feb. 2003, Swindle P.W., *et al.*, Urologic Clinics of North America. 30(2):377-401, May 2003.

[2] Wahlby C., *et al.*, Analytical Cellular Pathology 24, 101-111, 2002.

[3] Street W.N., "Xcvt: A System for Remote Cytological Diagnosis and Prognosis of Breast Cancer," In Soft Computing Techniques in Breast Cancer Prognosis and Diagnosis, L.C. Jain (ed.), CRC Press, 1999

20

[4] Gleason D.F., "The Veteran's Administration Cooperative Urologic Research Group: Histologic Grading and Clinical Staging of Prostatic Carcinoma," In Urologic Pathology: The Prostate, Tannenbaum M. (ed.), 171-198, Lea and Febiger, Philadelphia, 1977.

[5] Cristianni *et al.*, An Introduction to Support Vector Machines, Cambridge University Press (2000).

30

[6] Hastie, The Elements of Statistical Learning, Springer (2001).

[7] F.E. Harrell *et al.*, "Evaluating the yield of medical tests," JAMA, 247(18):2543-2546, 1982.

[8] Bishop, C., Neural Networks for Pattern Recognition, Oxford University Press (1995).

40

[9] Fausett, L., Fundamentals of Neural Networks, New York, Prentice Hall (1994).

[10] Definiens Cellenger Architecture: A Technical Review, April 2004.

【0323】

【化 3】

- [11] Baatz M. and Schäpe A., "Multiresolution Segmentation – An Optimization Approach for High Quality Multi-scale Image Segmentation," In Angewandte Geographische Informationsverarbeitung XII, Strobl, J., Blaschke, T., Griesebner, G. (eds.), Wichmann- Verlag, Heidelberg, 12-23, 2000.
- [12] Fukunaga K., Introduction to Statistical Pattern Recognition, 2nd Edition, Boston: Academic Press, 1990. 10
- [13] Duda R.O. *et al.*, Pattern Classification, 2nd Edition, John Wiley & Sons Inc., 2001.
- [14] Holmberg L. *et al.*, A randomized trial comparing radical prostatectomy with watchful waiting in early prostate cancer, *N. Engl. M. Med.*, 347:781-789 (2002).
- [15] Pound CR *et al.*, Natural history of progression after PSA elevation following radical prostatectomy, *JAMA* 1999, 281:1591-1597. 20
- [16] Kumar-Sinha C. *et al.*, Molecular markers to identify patients at risk for recurrence after primary treatment for prostate cancer, *Urology* 2003; 62 Suppl. 1:19-35.
- [17] Cox D.R., "Regression Models and Life Tables," Journal of the Royal Statistical Society, B 34, 187-220, 1972. 30
- [18] Harrell F.E., Regression Modeling Strategies, Springer-Verlag 2001.
- [19] Tuxhorn *et al.*, "Reactive Stroma in Human Prostate Cancer: Induction of Myofibroblast Phenotype and Extracellular Matrix Remodeling" *Clinical Cancer Research* 2912 Vol. 8, 2912–2923, September 2002.
- [20] Kattan *et al.*, "Postoperative Nomogram for Disease Recurrence After Radical Prostatectomy for Prostate Cancer," *Journal of Clinical Oncology*, Vol. 17, No. 5 (May), 1999: pp 1499-1507. 40

【 0 3 2 4 】

【化4】

- [21] C. Vonesch, F. Aquet, J.L. Vonesch and M. Unser, "The colored revolution of bioimaging," *IEEE Signal Proc. Mag.*, vol. 23, no. 3, pp. 20-31, May 2006.
- [22] A. Krtolica, C. O. de Solorzano, S. Lockett and J. Campisi, "Quantification of epithelial cells in coculture with fibroblast by fluorescence image analysis," *Cytometry*, vol. 49, pp. 73-82, 2002.
- [23] A. Gordon, A. Colman-Lerner, T. E. Chin, K. R. Benjamin, R. C. Yu, and R. Brent, "Single-cell quantification of molecules and rates using open-source microscope-based cytometry," *Nature Methods*, vol. 4, pp. 175-181, 2007. 10
- [24] R. Camp, G. G. Chung, and D. L. Rimm, "Automated subcellular localization and quantification of protein expression in tissue microarrays," *Nature Medicine*, vol. 8, pp. 1323-1327, 2002.
- [25] J. Y. Rao, D. Seligson, and G. P. Hemstreet, "Protein expression analysis using quantitative fluorescence image analysis on tissue microarray slides," *BioTechniques*, vol. 32, pp. 924-932, 2002.
- [26] Definiens Understanding Images, Developer Version 6, 2007.
<http://www.definiens.com>. 20
- [27] <http://www.cri-inc.com>.
- [28] A. Tabesh, M. Teverovskiy, H.-Y Pang, V. P. Kumar, D. Verbel, A. Kotsianti, and O. Saidi, "Multifeature prostate cancer diagnosis and Gleason grading of histological images," *IEEE Trans. Medical Imag.*, vol. 26, pp. 1366-1378, 2007.
- [29] Giaccone G, Herbst RS, Manegold C, et al: Gefitinib in combination with gemcitabine and cisplatin in advanced non-small-cell lung cancer: a phase III trial--INTACT 1. *J Clin Oncol* 22:777-84, 2004.
- [30] Herbst RS, Giaccone G, Schiller JH, et al: Gefitinib in combination with paclitaxel and carboplatin in advanced non-small-cell lung cancer: a phase III trial--INTACT 2. *J Clin Oncol* 22:785-94, 2004. 30
- [31] Thatcher N, Chang A, Parikh P, et al. Gefitinib plus best supportive care in previously treated patients with refractory advanced non-small-cell lung cancer: results from a randomised, placebo-controlled, multicentre study (Iressa Survival Evaluation in Lung Cancer). *Lancet* 366:1527-37, 2005.
- [32] Lynch TJ, Bell DW, Sordella R, et al: Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* 350:2129-39, 2004.

【0325】

40

【化5】

- [33] Dudek AZ, Kmak KL, Koopmeiners J, et al: Skin rash and bronchoalveolar histology correlates with clinical benefit in patients treated with gefitinib as a therapy for previously treated advanced or metastatic non-small cell lung cancer. *Lung Cancer* 51:89-96, 2006.
- [34] Cappuzzo F, Hirsch FR, Rossi E, et al: Epidermal growth factor receptor gene and protein and gefitinib sensitivity in non-small-cell lung cancer. *J Natl Cancer Inst* 97:643-55, 2005.
- [35] Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden et al: Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol* 5:649-655, 1982. Cordon-Cardo C, Kotsianti A, Donovan M, et al: Improved prediction of PSA recurrence through systems pathology. *J Clin Oncol* 22:4591, 2004 (abstr). 10
- [36] Zubek V, Verbel D, Saidi O: Censored time trees for predicting time to PSA recurrence, in Proceedings of the Fourth International Conference on Machine Learning and Applications (ICMLA 2005). Washington, DC, IEEE Computer Society, 2005, pp 221-226.
- [37] Paez JG, Janne PA, Lee JC, et al: EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* 304:1497-500, 2004. 20
- [38] Bell DW, Gore I, Okimoto RA, et al: Inherited susceptibility to lung cancer may be associated with the T790M drug resistance mutation in EGFR. *Nat Genet* 37:1315-6, 2005.
- [39] Hirsch FR, Varella-Garcia M, Bunn PA, Jr., et al: Epidermal growth factor receptor in non-small-cell lung carcinomas: correlation between gene copy number and protein expression and impact on prognosis. *J Clin Oncol* 21:3798-807, 2003.
- [40] Chan SK, Gullick WJ, Hill ME: Mutations of the epidermal growth factor receptor in non-small cell lung cancer -- search and destroy. *Eur J Cancer* 42:17-23, 2006. 30
- [41] Riely GJ, Pao W, Pham D, et al: Clinical course of patients with non-small cell lung cancer and epidermal growth factor receptor exon 19 and exon 21 mutations treated with gefitinib or erlotinib. *Clin Cancer Res* 12:839-44, 2006.
- [42] Sequist LV, Bell DW, Lynch TJ, et al: Molecular predictors of response to epidermal growth factor receptor antagonists in non-small-cell lung cancer. *J Clin Oncol* 25:587-95, 2007.

【0326】

40

【化6】

- [43] Uchida A, Hirano S, Kitao H, et al: Activation of downstream epidermal growth factor receptor (EGFR) signaling provides gefitinib-resistance in cells carrying EGFR mutation. *Cancer Sci* 98:357-63, 2007.
- [44] Sugio K, Uramoto H, Ono K, et al: Mutations within the tyrosine kinase domain of EGFR gene specifically occur in lung adenocarcinoma patients with a low exposure of tobacco smoking. *Br J Cancer* 94:896-903, 2006.
- [45] Edelman MJ: An update on the role of epidermal growth factor receptor inhibitors in non-small cell lung cancer. *Semin Oncol* 32:S3-8, 2005. 10
- [46] Kris MG, Natale RB, Herbst RS, et al: Efficacy of gefitinib, an inhibitor of the epidermal growth factor receptor tyrosine kinase, in symptomatic patients with non-small cell lung cancer: a randomized trial. *Jama* 290:2149-58, 2003.
- [47] Al-Kuraya K, Siraj AK, Bavi P, et al: High epidermal growth factor receptor amplification rate but low mutation frequency in Middle East lung cancer population. *Hum Pathol* 37:453-7, 2006.
- [48] Hainsworth JD, Mainwaring MG, Thomas M, et al: Gefitinib in the treatment of advanced, refractory non-small-cell lung cancer: results in 124 patients. *Clin Lung Cancer* 4:347-55, 2003. 20
- [49] Su WP, Yang CH, Yu CJ, et al: Gefitinib treatment for non-small cell lung cancer -- a study including patients with poor performance status. *J Formos Med Assoc* 104:557-62, 2005.
- [50] Erman M, Grunenwald D, Penault-Llorca F, et al: Epidermal growth factor receptor, HER-2/neu and related pathways in lung adenocarcinomas with bronchioloalveolar features. *Lung Cancer* 47:315-23, 2005.
- [51] Massion PP, Taflan PM, Shyr Y, et al: Early involvement of the phosphatidylinositol 3-kinase/Akt pathway in lung cancer progression. *Am J Respir Crit Care Med* 170:1088-94, 2004. 30
- [52] Cappuzzo F, Magrini E, Ceresoli GL, et al: Akt phosphorylation and gefitinib efficacy in patients with advanced non-small-cell lung cancer. *J Natl Cancer Inst* 96:1133-41, 2004.
- [53] Han SW, Hwang PG, Chung DH, et al: Epidermal growth factor receptor (EGFR) downstream molecules as response predictive markers for gefitinib (Iressa, ZD1839) in chemotherapy-resistant non-small cell lung cancer. *Int J Cancer* 113:109-15, 2005.
- [54] Bailey L, Kris M, Wolf M, et al: Tumor EGFR membrane staining is not clinically relevant for predicting response in patients receiving gefitinib (Iressa, 40

【0327】

【化 7】

- ZD 1839) monotherapy for pretreated advanced non-small-cell lung cancer: IDEAL 1 and 2. Proc Am Assoc Cancer Res 44:1362, 2003 (abstr.).
- [55] Bailey L, Janas M, Schmidt K, et al: Evaluation of epidermal growth factor receptor (EGFR) as a predictive marker in patients with non-small-cell lung cancer (NSCLC) receiving first-line gefitinib combined with platinum-based chemotherapy. J Clin Oncol 22:7013, 2004 (abstr.).
- [56] Jin M, Inoue S, Umemura T, et al: Cyclin D1, p16 and retinoblastoma gene product expression as a predictor for prognosis in non-small cell lung cancer at stages I and II. Lung Cancer 34:207-18, 2001. 10
- [57] Keum JS, Kong G, Yang SC, et al: Cyclin D1 overexpression is an indicator of poor prognosis in resectable non-small cell lung cancer. Br J Cancer 81:127-32, 1999.
- [58] Kalish LH, Kwong RA, Cole IE, et al: Deregulated cyclin D1 expression is associated with decreased efficacy of the selective epidermal growth factor receptor tyrosine kinase inhibitor gefitinib in head and neck squamous cell carcinoma cell lines. Clin Cancer Res 10:7764-74, 2004.
- [59] Koomagi R, Volm M: Relationship between the expression of caspase-3 and the clinical outcome of patients with non-small cell lung cancer. Anticancer Res 20:493-6, 2000. 20
- [60] Takata T, Tanaka F, Yamada T, et al: Clinical significance of caspase-3 expression in pathologic-stage I, nonsmall-cell lung cancer. Int J Cancer 96 Suppl:54-60, 2001.
- [61] Mineo TC, Ambrogi V, Baldi A, et al: Prognostic impact of VEGF, CD31, CD34, and CD105 expression and tumour vessel invasion after radical surgery for IB-IIA non-small cell lung cancer. J Clin Pathol 57:591-7, 2004. 30

【 0 3 2 8 】

【 表 1 - 1 】

表1. 形態計測的フィーチャ
Script v1.0 (496 フィーチャ)

フィーチャ	
Background.MaxAreaPx1	
Background.MeanAreaPx1	
Background.MinAreaPx1	
Background.StdDevAreaPx1	
Background.SumAreaPx1	40
Cytoplasm.Objects	
Cytoplasm.ObjectsPct	
Cytoplasm.MaxAreaPx1	
Cytoplasm.MeanAreaPx1	

【 0 3 2 9 】

【表 1 - 2】

フィーチャ	
Cytoplasm.MinAreaPxl	
Cytoplasm.StdDevAreaPxl	
Cytoplasm.SumAreaPxl	
Cytoplasm.MaxAsymmetry	
Cytoplasm.MeanAsymmetry	
Cytoplasm.MinAsymmetry	
Cytoplasm.StdDevAsymmetry	10
Cytoplasm.MaxBorderlengthPxl	
Cytoplasm.MeanBorderlengthPxl	
Cytoplasm.MinBorderlengthPxl	
Cytoplasm.StdDevBorderlengthPxl	
Cytoplasm.SumBorderlengthPxl	
Cytoplasm.MaxBrightness	
Cytoplasm.MeanBrightness	
Cytoplasm.MinBrightness	
Cytoplasm.StdDevBrightness	
Cytoplasm.MaxCompactness	
Cytoplasm.MeanCompactness	20
Cytoplasm.MinCompactness	
Cytoplasm.StdDevCompactness	
Cytoplasm.MaxDensity	
Cytoplasm.MeanDensity	
Cytoplasm.MinDensity	
Cytoplasm.StdDevDensity	
Cytoplasm.MaxDiff.ofenclosing.encl	
Cytoplasm.MeanDiff.ofenclosing.encl	
Cytoplasm.MinDiff.ofenclosing.encl	
Cytoplasm.StdDevDiff.ofenclosing.en	30
Cytoplasm.MaxEllipticFit	
Cytoplasm.MeanEllipticFit	
Cytoplasm.MinEllipticFit	
Cytoplasm.StdDevEllipticFit	
Cytoplasm.MaxLengthPxl	
Cytoplasm.MeanLengthPxl	
Cytoplasm.MinLengthPxl	
Cytoplasm.StdDevLengthPxl	
Cytoplasm.SumLengthPxl	
Cytoplasm.MaxMax.Diff.	
Cytoplasm.MeanMax.Diff.	40
Cytoplasm.MinMax.Diff.	

【 0 3 3 0 】

【表 1 - 3】

フィーチャ	
Cytoplasm.StdDevMax.Diff.	
Cytoplasm.MaxMeanChannel1	
Cytoplasm.MeanMeanChannel1	
Cytoplasm.MinMeanChannel1	
Cytoplasm.StdDevMeanChannel1	
Cytoplasm.MaxMeanChannel2	
Cytoplasm.MeanMeanChannel2	
Cytoplasm.MinMeanChannel2	10
Cytoplasm.StdDevMeanChannel2	
Cytoplasm.MaxMeanChannel3	
Cytoplasm.MeanMeanChannel3	
Cytoplasm.MinMeanChannel3	
Cytoplasm.StdDevMeanChannel3	
Cytoplasm.MaxRadiusoflargestenclose	
Cytoplasm.MeanRadiusoflargestenclos	
Cytoplasm.MinRadiusoflargestenclose	
Cytoplasm.StdDevRadiusoflargestenclo	
Cytoplasm.MaxRadiusofsmallestenclos	20
Cytoplasm.MeanRadiusofsmallestenclo	
Cytoplasm.MinRadiusofsmallestenclos	
Cytoplasm.StdDevRadiusofsmallestenc	
Cytoplasm.MaxStdevChannel1	
Cytoplasm.MeanStdevChannel1	
Cytoplasm.MinStdevChannel1	
Cytoplasm.StdDevStdevChannel1	
Cytoplasm.MaxStdevChannel2	
Cytoplasm.MeanStdevChannel2	
Cytoplasm.MinStdevChannel2	
Cytoplasm.StdDevStdevChannel2	30
Cytoplasm.MaxStdevChannel3	
Cytoplasm.MeanStdevChannel3	
Cytoplasm.MinStdevChannel3	
Cytoplasm.StdDevStdevChannel3	
Cytoplasm.MaxWidthPx1	
Cytoplasm.MeanWidthPx1	
Cytoplasm.MinWidthPx1	
Cytoplasm.StdDevWidthPx1	
Epithelial.Nuclei.Objects	
Epithelial.Nuclei.ObjectsPct	40
Epithelial.Nuclei.MaxAreaPx1	

【 0 3 3 1 】

【表 1 - 4】

 フィーチャ

Epithelial.Nuclei.MeanAreaPxI	
Epithelial.Nuclei.MinAreaPxI	
Epithelial.Nuclei.StdDevAreaPxI	
Epithelial.Nuclei.SumAreaPxI	
Epithelial.Nuclei.MaxAsymmetry	
Epithelial.Nuclei.MeanAsymmetry	
Epithelial.Nuclei.MinAsymmetry	10
Epithelial.Nuclei.StdDevAsymmetry	
Epithelial.Nuclei.MaxBorderlengthPx	
Epithelial.Nuclei.MeanBorderlengthP	
Epithelial.Nuclei.MinBorderlengthPx	
Epithelial.Nuclei.StdDevBorderlengt	
Epithelial.Nuclei.SumBorderlengthPx	
Epithelial.Nuclei.MaxBrightness	
Epithelial.Nuclei.MeanBrightness	
Epithelial.Nuclei.MinBrightness	
Epithelial.Nuclei.StdDevBrightness	
Epithelial.Nuclei.MaxCompactness	20
Epithelial.Nuclei.MeanCompactness	
Epithelial.Nuclei.MinCompactness	
Epithelial.Nuclei.StdDevCompactness	
Epithelial.Nuclei.MaxDensity	
Epithelial.Nuclei.MeanDensity	
Epithelial.Nuclei.MinDensity	
Epithelial.Nuclei.StdDevDensity	
Epithelial.Nuclei.MaxDiff.ofenclosi	
Epithelial.Nuclei.MeanDiff.ofenclos	
Epithelial.Nuclei.MinDiff.ofenclosi	30
Epithelial.Nuclei.StdDevDiff.ofencl	
Epithelial.Nuclei.MaxEllipticFit	
Epithelial.Nuclei.MeanEllipticFit	
Epithelial.Nuclei.MinEllipticFit	
Epithelial.Nuclei.StdDevEllipticFit	
Epithelial.Nuclei.MaxLengthPxI	
Epithelial.Nuclei.MeanLengthPxI	
Epithelial.Nuclei.MinLengthPxI	
Epithelial.Nuclei.StdDevLengthPxI	
Epithelial.Nuclei.SumLengthPxI	
Epithelial.Nuclei.MaxMax.Diff.	40
Epithelial.Nuclei.MeanMax.Diff.	

【 0 3 3 2 】

【表 1 - 5】

フィーチャ

Epithelial.Nuclei.MinMax.Diff.	
Epithelial.Nuclei.StdDevMax.Diff.	
Epithelial.Nuclei.MaxMeanChannel1	
Epithelial.Nuclei.MeanMeanChannel1	
Epithelial.Nuclei.MinMeanChannel1	
Epithelial.Nuclei.StdDevMeanChannel1	
Epithelial.Nuclei.MaxMeanChannel2	10
Epithelial.Nuclei.MeanMeanChannel2	
Epithelial.Nuclei.MinMeanChannel2	
Epithelial.Nuclei.StdDevMeanChannel2	
Epithelial.Nuclei.MaxMeanChannel3	
Epithelial.Nuclei.MeanMeanChannel3	
Epithelial.Nuclei.MinMeanChannel3	
Epithelial.Nuclei.StdDevMeanChannel3	
Epithelial.Nuclei.MaxRadiusoflarges	
Epithelial.Nuclei.MeanRadiusoflarge	
Epithelial.Nuclei.MinRadiusoflarges	
Epithelial.Nuclei.StdDevRadiusoflar	20
Epithelial.Nuclei.MaxRadiusofsmalle	
Epithelial.Nuclei.MeanRadiusofsmall	
Epithelial.Nuclei.MinRadiusofsmalle	
Epithelial.Nuclei.StdDevRadiusofsma	
Epithelial.Nuclei.MaxStdevChannel1	
Epithelial.Nuclei.MeanStdevChannel1	
Epithelial.Nuclei.MinStdevChannel1	
Epithelial.Nuclei.StdDevStdevChanne	
Epithelial.Nuclei.MaxStdevChannel2	
Epithelial.Nuclei.MeanStdevChannel2	30
Epithelial.Nuclei.MinStdevChannel2	
Epithelial.Nuclei.StdDevStdevChanne	
Epithelial.Nuclei.MaxStdevChannel3	
Epithelial.Nuclei.MeanStdevChannel3	
Epithelial.Nuclei.MinStdevChannel3	
Epithelial.Nuclei.StdDevStdevChanne	
Epithelial.Nuclei.MaxWidthPx1	
Epithelial.Nuclei.MeanWidthPx1	
Epithelial.Nuclei.MinWidthPx1	
Epithelial.Nuclei.StdDevWidthPx1	
Lumen.Objects	40
Lumen.ObjectsPct	

【 0 3 3 3 】

【表 1 - 6】

フィーチャ	
Lumen.MaxAreaPxI	
Lumen.MeanAreaPxI	
Lumen.MinAreaPxI	
Lumen.StdDevAreaPxI	
Lumen.SumAreaPxI	
Lumen.MaxAsymmetry	
Lumen.MeanAsymmetry	10
Lumen.MinAsymmetry	
Lumen.StdDevAsymmetry	
Lumen.MaxBorderlengthPxI	
Lumen.MeanBorderlengthPxI	
Lumen.MinBorderlengthPxI	
Lumen.StdDevBorderlengthPxI	
Lumen.SumBorderlengthPxI	
Lumen.MaxBrightness	
Lumen.MeanBrightness	
Lumen.MinBrightness	
Lumen.StdDevBrightness	20
Lumen.MaxCompactness	
Lumen.MeanCompactness	
Lumen.MinCompactness	
Lumen.StdDevCompactness	
Lumen.MaxDensity	
Lumen.MeanDensity	
Lumen.MinDensity	
Lumen.StdDevDensity	
Lumen.MaxDiff.ofenclosing.enclosede	
Lumen.MeanDiff.ofenclosing.enclosed	30
Lumen.MinDiff.ofenclosing.enclosede	
Lumen.StdDevDiff.ofenclosing.enclos	
Lumen.MaxEllipticFit	
Lumen.MeanEllipticFit	
Lumen.MinEllipticFit	
Lumen.StdDevEllipticFit	
Lumen.MaxLengthPxI	
Lumen.MeanLengthPxI	
Lumen.MinLengthPxI	
Lumen.StdDevLengthPxI	40
Lumen.SumLengthPxI	
Lumen.MaxMax.Diff.	

【 0 3 3 4 】

【表 1 - 7】

フィーチャ	
Lumen.MeanMax.Diff.	
Lumen.MinMax.Diff.	
Lumen.StdDevMax.Diff.	
Lumen.MaxMeanChannel1	
Lumen.MeanMeanChannel1	
Lumen.MinMeanChannel1	
Lumen.StdDevMeanChannel1	10
Lumen.MaxMeanChannel2	
Lumen.MeanMeanChannel2	
Lumen.MinMeanChannel2	
Lumen.StdDevMeanChannel2	
Lumen.MaxMeanChannel3	
Lumen.MeanMeanChannel3	
Lumen.MinMeanChannel3	
Lumen.StdDevMeanChannel3	
Lumen.MaxRadiusoflargestenclosedell	
Lumen.MeanRadiusoflargestenclosedel	
Lumen.MinRadiusoflargestenclosedell	20
Lumen.StdDevRadiusoflargestenclosed	
Lumen.MaxRadiusofsmallestenclosinge	
Lumen.MeanRadiusofsmallestenclosing	
Lumen.MinRadiusofsmallestenclosinge	
Lumen.StdDevRadiusofsmallestenclosi	
Lumen.MaxStdevChannel1	
Lumen.MeanStdevChannel1	
Lumen.MinStdevChannel1	
Lumen.StdDevStdevChannel1	
Lumen.MaxStdevChannel2	
Lumen.MeanStdevChannel2	
Lumen.MinStdevChannel2	30
Lumen.StdDevStdevChannel2	
Lumen.MaxStdevChannel3	
Lumen.MeanStdevChannel3	
Lumen.MinStdevChannel3	
Lumen.StdDevStdevChannel3	
Lumen.MaxWidthPx1	
Lumen.MeanWidthPx1	
Lumen.MinWidthPx1	
Lumen.StdDevWidthPx1	
Red.Blood.Cell.Objects	
Red.Blood.Cell.ObjectsPct	
Red.Blood.Cell.MaxAreaPx1	40
Red.Blood.Cell.MeanAreaPx1	

【 0 3 3 5 】

【表 1 - 8】

フィーチャ	
Red.Blood.Cell.MinAreaPx1	
Red.Blood.Cell.StdDevAreaPx1	
Red.Blood.Cell.SumAreaPx1	
Red.Blood.Cell.MaxAsymmetry	
Red.Blood.Cell.MeanAsymmetry	
Red.Blood.Cell.MinAsymmetry	
Red.Blood.Cell.StdDevAsymmetry	
Red.Blood.Cell.MaxBorderlengthPx1	10
Red.Blood.Cell.MeanBorderlengthPx1	
Red.Blood.Cell.MinBorderlengthPx1	
Red.Blood.Cell.StdDevBorderlengthPx	
Red.Blood.Cell.SumBorderlengthPx1	
Red.Blood.Cell.MaxBrightness	
Red.Blood.Cell.MeanBrightness	
Red.Blood.Cell.MinBrightness	
Red.Blood.Cell.StdDevBrightness	
Red.Blood.Cell.MaxCompactness	
Red.Blood.Cell.MeanCompactness	20
Red.Blood.Cell.MinCompactness	
Red.Blood.Cell.StdDevCompactness	
Red.Blood.Cell.MaxDensity	
Red.Blood.Cell.MeanDensity	
Red.Blood.Cell.MinDensity	
Red.Blood.Cell.StdDevDensity	
Red.Blood.Cell.MaxDiff.ofenclosing.	
Red.Blood.Cell.MeanDiff.ofenclosing	
Red.Blood.Cell.MinDiff.ofenclosing.	
Red.Blood.Cell.StdDevDiff.ofenclosi	
Red.Blood.Cell.MaxEllipticFit	30
Red.Blood.Cell.MeanEllipticFit	
Red.Blood.Cell.MinEllipticFit	
Red.Blood.Cell.StdDevEllipticFit	
Red.Blood.Cell.MaxLengthPx1	
Red.Blood.Cell.MeanLengthPx1	
Red.Blood.Cell.MinLengthPx1	
Red.Blood.Cell.StdDevLengthPx1	
Red.Blood.Cell.SumLengthPx1	
Red.Blood.Cell.MaxMax.Diff.	
Red.Blood.Cell.MeanMax.Diff.	40
Red.Blood.Cell.MinMax.Diff.	

【 0 3 3 6 】

【表 1 - 9】

フィーチャ

Red.Blood.Cell.StdDevMax.Diff.	
Red.Blood.Cell.MaxMeanChannel1	
Red.Blood.Cell.MeanMeanChannel1	
Red.Blood.Cell.MinMeanChannel1	
Red.Blood.Cell.StdDevMeanChannel1	
Red.Blood.Cell.MaxMeanChannel2	
Red.Blood.Cell.MeanMeanChannel2	
Red.Blood.Cell.MinMeanChannel2	10
Red.Blood.Cell.StdDevMeanChannel2	
Red.Blood.Cell.MaxMeanChannel3	
Red.Blood.Cell.MeanMeanChannel3	
Red.Blood.Cell.MinMeanChannel3	
Red.Blood.Cell.StdDevMeanChannel3	
Red.Blood.Cell.MaxRadiusoflargesten	
Red.Blood.Cell.MeanRadiusoflargeste	
Red.Blood.Cell.MinRadiusoflargesten	
Red.Blood.Cell.StdDevRadiusoflarges	
Red.Blood.Cell.MaxRadiusofsmalleste	20
Red.Blood.Cell.MeanRadiusofsmallest	
Red.Blood.Cell.MinRadiusofsmalleste	
Red.Blood.Cell.StdDevRadiusofsmalle	
Red.Blood.Cell.MaxStdevChannel1	
Red.Blood.Cell.MeanStdevChannel1	
Red.Blood.Cell.MinStdevChannel1	
Red.Blood.Cell.StdDevStdevChannel1	
Red.Blood.Cell.MaxStdevChannel2	
Red.Blood.Cell.MeanStdevChannel2	
Red.Blood.Cell.MinStdevChannel2	
Red.Blood.Cell.StdDevStdevChannel2	30
Red.Blood.Cell.MaxStdevChannel3	
Red.Blood.Cell.MeanStdevChannel3	
Red.Blood.Cell.MinStdevChannel3	
Red.Blood.Cell.StdDevStdevChannel3	
Red.Blood.Cell.MaxWidthPx1	
Red.Blood.Cell.MeanWidthPx1	
Red.Blood.Cell.MinWidthPx1	
Red.Blood.Cell.StdDevWidthPx1	
Stroma.Objects	
Stroma.ObjectsPct	40
Stroma.MaxAreaPx1	

【 0 3 3 7 】

【表 1 - 1 0】

フィーチャ	
Stroma.MeanAreaPx1	
Stroma.MinAreaPx1	
Stroma.StdDevAreaPx1	
Stroma.SumAreaPx1	
Stroma.MaxAsymmetry	
Stroma.MeanAsymmetry	
Stroma.MinAsymmetry	10
Stroma.StdDevAsymmetry	
Stroma.MaxBorderlengthPx1	
Stroma.MeanBorderlengthPx1	
Stroma.MinBorderlengthPx1	
Stroma.StdDevBorderlengthPx1	
Stroma.SumBorderlengthPx1	
Stroma.MaxBrightness	
Stroma.MeanBrightness	
Stroma.MinBrightness	
Stroma.StdDevBrightness	
Stroma.MaxCompactness	20
Stroma.MeanCompactness	
Stroma.MinCompactness	
Stroma.StdDevCompactness	
Stroma.MaxDensity	
Stroma.MeanDensity	
Stroma.MinDensity	
Stroma.StdDevDensity	
Stroma.MaxDiff.ofenclosing.enclosed	
Stroma.MeanDiff.ofenclosing.encluse	
Stroma.MinDiff.ofenclosing.enclosed	30
Stroma.StdDevDiff.ofenclosing.enclou	
Stroma.MaxEllipticFit	
Stroma.MeanEllipticFit	
Stroma.MinEllipticFit	
Stroma.StdDevEllipticFit	
Stroma.MaxLengthPx1	
Stroma.MeanLengthPx1	
Stroma.MinLengthPx1	
Stroma.StdDevLengthPx1	
Stroma.SumLengthPx1	
Stroma.MaxMax.Diff.	40
Stroma.MeanMax.Diff.	

【 0 3 3 8 】

【表 1 - 1 1】

フイーチャ	
Stroma.MinMax.Diff.	
Stroma.StdDevMax.Diff.	
Stroma.MaxMeanChannel1	
Stroma.MeanMeanChannel1	
Stroma.MinMeanChannel1	
Stroma.StdDevMeanChannel1	
Stroma.MaxMeanChannel2	
Stroma.MeanMeanChannel2	10
Stroma.MinMeanChannel2	
Stroma.StdDevMeanChannel2	
Stroma.MaxMeanChannel3	
Stroma.MeanMeanChannel3	
Stroma.MinMeanChannel3	
Stroma.StdDevMeanChannel3	
Stroma.MaxRadiusoflargestenclosedel	
Stroma.MeanRadiusoflargestenclosede	
Stroma.MinRadiusoflargestenclosedel	
Stroma.StdDevRadiusoflargestenclose	20
Stroma.MaxRadiusofsmallestenclosing	
Stroma.MeanRadiusofsmallestenclosin	
Stroma.MinRadiusofsmallestenclosing	
Stroma.StdDevRadiusofsmallestenclos	
Stroma.MaxStdevChannel1	
Stroma.MeanStdevChannel1	
Stroma.MinStdevChannel1	
Stroma.StdDevStdevChannel1	
Stroma.MaxStdevChannel2	
Stroma.MeanStdevChannel2	
Stroma.MinStdevChannel2	30
Stroma.StdDevStdevChannel2	
Stroma.MaxStdevChannel3	
Stroma.MeanStdevChannel3	
Stroma.MinStdevChannel3	
Stroma.StdDevStdevChannel3	
Stroma.MaxWidthPx1	
Stroma.MeanWidthPx1	
Stroma.MinWidthPx1	
Stroma.StdDevWidthPx1	
Stroma.Nuclei.Objects	40
Stroma.Nuclei.ObjectsPct	

【 0 3 3 9 】

【表 1 - 1 2】

フィーチャ

Stroma.Nuclei.MaxAreaPx1	
Stroma.Nuclei.MeanAreaPx1	
Stroma.Nuclei.MinAreaPx1	
Stroma.Nuclei.StdDevAreaPx1	
Stroma.Nuclei.SumAreaPx1	
Stroma.Nuclei.MaxAsymmetry	
Stroma.Nuclei.MeanAsymmetry	10
Stroma.Nuclei.MinAsymmetry	
Stroma.Nuclei.StdDevAsymmetry	
Stroma.Nuclei.MaxBorderlengthPx1	
Stroma.Nuclei.MeanBorderlengthPx1	
Stroma.Nuclei.MinBorderlengthPx1	
Stroma.Nuclei.StdDevBorderlengthPx1	
Stroma.Nuclei.SumBorderlengthPx1	
Stroma.Nuclei.MaxBrightness	
Stroma.Nuclei.MeanBrightness	
Stroma.Nuclei.MinBrightness	
Stroma.Nuclei.StdDevBrightness	20
Stroma.Nuclei.MaxCompactness	
Stroma.Nuclei.MeanCompactness	
Stroma.Nuclei.MinCompactness	
Stroma.Nuclei.StdDevCompactness	
Stroma.Nuclei.MaxDensity	
Stroma.Nuclei.MeanDensity	
Stroma.Nuclei.MinDensity	
Stroma.Nuclei.StdDevDensity	
Stroma.Nuclei.MaxDiff.ofenclosing.e	
Stroma.Nuclei.MeanDiff.ofenclosing.e	30
Stroma.Nuclei.MinDiff.ofenclosing.e	
Stroma.Nuclei.StdDevDiff.ofenclosin	
Stroma.Nuclei.MaxEllipticFit	
Stroma.Nuclei.MeanEllipticFit	
Stroma.Nuclei.MinEllipticFit	
Stroma.Nuclei.StdDevEllipticFit	
Stroma.Nuclei.MaxLengthPx1	
Stroma.Nuclei.MeanLengthPx1	
Stroma.Nuclei.MinLengthPx1	
Stroma.Nuclei.StdDevLengthPx1	
Stroma.Nuclei.SumLengthPx1	40
Stroma.Nuclei.MaxMax.Diff.	

【 0 3 4 0 】

【表 1 - 1 3】

フィーチャ	
Stroma.Nuclei.MeanMax.Diff.	
Stroma.Nuclei.MinMax.Diff.	
Stroma.Nuclei.StdDevMax.Diff.	
Stroma.Nuclei.MaxMeanChannel1	
Stroma.Nuclei.MeanMeanChannel1	
Stroma.Nuclei.MinMeanChannel1	
Stroma.Nuclei.StdDevMeanChannel1	10
Stroma.Nuclei.MaxMeanChannel2	
Stroma.Nuclei.MeanMeanChannel2	
Stroma.Nuclei.MinMeanChannel2	
Stroma.Nuclei.StdDevMeanChannel2	
Stroma.Nuclei.MaxMeanChannel3	
Stroma.Nuclei.MeanMeanChannel3	
Stroma.Nuclei.MinMeanChannel3	
Stroma.Nuclei.StdDevMeanChannel3	
Stroma.Nuclei.MaxRadiusoflargestenc	
Stroma.Nuclei.MeanRadiusoflargestenc	
Stroma.Nuclei.MinRadiusoflargestenc	
Stroma.Nuclei.StdDevRadiusoflargest	
Stroma.Nuclei.MaxRadiusofsmallestenc	20
Stroma.Nuclei.MeanRadiusofsmallestenc	
Stroma.Nuclei.MinRadiusofsmallestenc	
Stroma.Nuclei.StdDevRadiusofsmalles	
Stroma.Nuclei.MaxStdevChannel1	
Stroma.Nuclei.MeanStdevChannel1	
Stroma.Nuclei.MinStdevChannel1	
Stroma.Nuclei.StdDevStdevChannel1	
Stroma.Nuclei.MaxStdevChannel2	
Stroma.Nuclei.MeanStdevChannel2	
Stroma.Nuclei.MinStdevChannel2	
Stroma.Nuclei.StdDevStdevChannel2	
Stroma.Nuclei.MaxStdevChannel3	
Stroma.Nuclei.MeanStdevChannel3	30
Stroma.Nuclei.MinStdevChannel3	
Stroma.Nuclei.StdDevStdevChannel3	
Stroma.Nuclei.MaxWidthPx1	
Stroma.Nuclei.MeanWidthPx1	
Stroma.Nuclei.MinWidthPx1	
Stroma.Nuclei.StdDevWidthPx1	
C2EN	
EN2SN	
L2Core	
C2L	
CEN2L	

【 0 3 4 1 】

40

【表 2 - 1】

表2. 形態計測的フィーチャ
Script v2. 0 (350フィーチャ)

フィーチャ	
Artifact Mean Area Pxl	
Artifact StdDev Area Pxl	
Artifact Mean Asymmetry	
Artifact StdDev Asymmetry	
Artifact Mean Border index	
Artifact StdDev Border index	
Artifact Mean Border length Pxl	10
Artifact StdDev Border length Pxl	
Artifact Mean Brightness	
Artifact StdDev Brightness	
Artifact Mean Compactness	
Artifact StdDev Compactness	
Artifact Mean Density	
Artifact StdDev Density	
Artifact Mean Diff. of enclosing/enclosed ellipse	
Artifact StdDev Diff. of enclosing/enclosed ellipse	
Artifact Mean Elliptic Fit	
Artifact StdDev Elliptic Fit	
Artifact Mean Length Pxl	20
Artifact StdDev Length Pxl	
Artifact Mean Length/width	
Artifact StdDev Length/width	
Artifact Mean Main direction	
Artifact StdDev Main direction	
Artifact Mean Max.Diff.	
Artifact StdDev Max.Diff.	
Artifact Mean Mean Channel 1	
Artifact StdDev Mean Channel 1	
Artifact Mean Mean Channel 2	
Artifact StdDev Mean Channel 2	
Artifact Mean Mean Channel 3	30
Artifact StdDev Mean Channel 3	
Artifact Mean Radius of largest enclosed ellipse	
Artifact StdDev Radius of largest enclosed ellipse	
Artifact Mean Radius of smallest enclosing ellipse	
Artifact StdDev Radius of smallest enclosing ellipse	
Artifact Mean Rectangular Fit	
Artifact StdDev Rectangular Fit	
Artifact Mean Shape index	
Artifact StdDev Shape index	
Artifact Mean Stddev Channel 1	
Artifact StdDev Stddev Channel 1	
Artifact Mean Stddev Channel 2	40
Artifact StdDev Stddev Channel 2	
Artifact Mean Stddev Channel 3	
Artifact StdDev Stddev Channel 3	

【 0 3 4 2 】

【表 2 - 2】

フィーチャ	
Artifact Mean Width Pxl	
Artifact StdDev Width Pxl	
Cytoplasm Mean Area Pxl	
Cytoplasm StdDev Area Pxl	
Cytoplasm Mean Asymmetry	
Cytoplasm StdDev Asymmetry	
Cytoplasm Mean Border index	
Cytoplasm StdDev Border index	
Cytoplasm Mean Border length Pxl	10
Cytoplasm StdDev Border length Pxl	
Cytoplasm Mean Brightness	
Cytoplasm StdDev Brightness	
Cytoplasm Mean Compactness	
Cytoplasm StdDev Compactness	
Cytoplasm Mean Density	
Cytoplasm StdDev Density	
Cytoplasm Mean Diff. of enclosing/enclosed ellipse	
Cytoplasm StdDev Diff. of enclosing/enclosed ellipse	
Cytoplasm Mean Elliptic Fit	
Cytoplasm StdDev Elliptic Fit	
Cytoplasm Mean Length Pxl	20
Cytoplasm StdDev Length Pxl	
Cytoplasm Mean Length/width	
Cytoplasm StdDev Length/width	
Cytoplasm Mean Main direction	
Cytoplasm StdDev Main direction	
Cytoplasm Mean Max.Diff.	
Cytoplasm StdDev Max.Diff.	
Cytoplasm Mean Mean Channel 1	
Cytoplasm StdDev Mean Channel 1	
Cytoplasm Mean Mean Channel 2	
Cytoplasm StdDev Mean Channel 2	
Cytoplasm Mean Mean Channel 3	
Cytoplasm StdDev Mean Channel 3	30
Cytoplasm Mean Radius of largest enclosed ellipse	
Cytoplasm StdDev Radius of largest enclosed ellipse	
Cytoplasm Mean Radius of smallest enclosing ellipse	
Cytoplasm StdDev Radius of smallest enclosing ellipse	
Cytoplasm Mean Rectangular Fit	
Cytoplasm StdDev Rectangular Fit	
Cytoplasm Mean Shape index	
Cytoplasm StdDev Shape index	
Cytoplasm Mean Stddev Channel 1	
Cytoplasm StdDev Stddev Channel 1	
Cytoplasm Mean Stddev Channel 2	
Cytoplasm StdDev Stddev Channel 2	40

【 0 3 4 3 】

【表 2 - 3】

フィーチャ

Cytoplasm Mean Stddev Channel 3	
Cytoplasm StdDev Stddev Channel 3	
Cytoplasm Mean Width Pxl	
Cytoplasm StdDev Width Pxl	
Epithelial Nuclei Mean Area Pxl	
Epithelial Nuclei StdDev Area Pxl	
Epithelial Nuclei Mean Asymmetry	
Epithelial Nuclei StdDev Asymmetry	
Epithelial Nuclei Mean Border index	10
Epithelial Nuclei StdDev Border index	
Epithelial Nuclei Mean Border length Pxl	
Epithelial Nuclei StdDev Border length Pxl	
Epithelial Nuclei Mean Brightness	
Epithelial Nuclei StdDev Brightness	
Epithelial Nuclei Mean Compactness	
Epithelial Nuclei StdDev Compactness	
Epithelial Nuclei Mean Density	
Epithelial Nuclei StdDev Density	
Epithelial Nuclei Mean Diff. of enclosing/enclosed ellipse	
Epithelial Nuclei StdDev Diff. of enclosing/enclosed ellipse	
Epithelial Nuclei Mean Elliptic Fit	20
Epithelial Nuclei StdDev Elliptic Fit	
Epithelial Nuclei Mean Length Pxl	
Epithelial Nuclei StdDev Length Pxl	
Epithelial Nuclei Mean Length/width	
Epithelial Nuclei StdDev Length/width	
Epithelial Nuclei Mean Main direction	
Epithelial Nuclei StdDev Main direction	
Epithelial Nuclei Mean Max.Diff.	
Epithelial Nuclei StdDev Max.Diff.	
Epithelial Nuclei Mean Mean Channel 1	
Epithelial Nuclei StdDev Mean Channel 1	
Epithelial Nuclei Mean Mean Channel 2	30
Epithelial Nuclei StdDev Mean Channel 2	
Epithelial Nuclei Mean Mean Channel 3	
Epithelial Nuclei StdDev Mean Channel 3	
Epithelial Nuclei Mean Radius of largest enclosed ellipse	
Epithelial Nuclei StdDev Radius of largest enclosed ellipse	
Epithelial Nuclei Mean Radius of smallest enclosing ellipse	
Epithelial Nuclei StdDev Radius of smallest enclosing ellipse	
Epithelial Nuclei Mean Rectangular Fit	
Epithelial Nuclei StdDev Rectangular Fit	
Epithelial Nuclei Mean Shape index	
Epithelial Nuclei StdDev Shape index	
Epithelial Nuclei Mean Stddev Channel 1	40
Epithelial Nuclei StdDev Stddev Channel 1	

【 0 3 4 4 】

【表 2 - 4】

フィーチャ	
Epithelial Nuclei Mean Stddev Channel 2	
Epithelial Nuclei StdDev Stddev Channel 2	
Epithelial Nuclei Mean Stddev Channel 3	
Epithelial Nuclei StdDev Stddev Channel 3	
Epithelial Nuclei Mean Width Pxl	
Epithelial Nuclei StdDev Width Pxl	
Lumen Mean Area Pxl	
Lumen StdDev Area Pxl	
Lumen Mean Asymmetry	10
Lumen StdDev Asymmetry	
Lumen Mean Border index	
Lumen StdDev Border index	
Lumen Mean Border length Pxl	
Lumen StdDev Border length Pxl	
Lumen Mean Brightness	
Lumen StdDev Brightness	
Lumen Mean Compactness	
Lumen StdDev Compactness	
Lumen Mean Density	
Lumen StdDev Density	
Lumen Mean Diff. of enclosing/enclosed ellipse	20
Lumen StdDev Diff. of enclosing/enclosed ellipse	
Lumen Mean Elliptic Fit	
Lumen StdDev Elliptic Fit	
Lumen Mean Length Pxl	
Lumen StdDev Length Pxl	
Lumen Mean Length/width	
Lumen StdDev Length/width	
Lumen Mean Main direction	
Lumen StdDev Main direction	
Lumen Mean Max.Diff.	
Lumen StdDev Max.Diff.	
Lumen Mean Mean Channel 1	
Lumen StdDev Mean Channel 1	30
Lumen Mean Mean Channel 2	
Lumen StdDev Mean Channel 2	
Lumen Mean Mean Channel 3	
Lumen StdDev Mean Channel 3	
Lumen Mean Radius of largest enclosed ellipse	
Lumen StdDev Radius of largest enclosed ellipse	
Lumen Mean Radius of smallest enclosing ellipse	
Lumen StdDev Radius of smallest enclosing ellipse	
Lumen Mean Rectangular Fit	
Lumen StdDev Rectangular Fit	
Lumen Mean Shape index	
Lumen StdDev Shape index	40

【 0 3 4 5 】

【表 2 - 5】

フィーチャ	
Lumen Mean Stddev Channel 1	
Lumen StdDev Stddev Channel 1	
Lumen Mean Stddev Channel 2	
Lumen StdDev Stddev Channel 2	
Lumen Mean Stddev Channel 3	
Lumen StdDev Stddev Channel 3	
Lumen Mean Width Pxl	
Lumen StdDev Width Pxl	
Stroma Mean Area Pxl	10
Stroma StdDev Area Pxl	
Stroma Mean Asymmetry	
Stroma StdDev Asymmetry	
Stroma Mean Border index	
Stroma StdDev Border index	
Stroma Mean Border length Pxl	
Stroma StdDev Border length Pxl	
Stroma Mean Brightness	
Stroma StdDev Brightness	
Stroma Mean Compactness	
Stroma StdDev Compactness	
Stroma Mean Density	20
Stroma StdDev Density	
Stroma Mean Diff. of enclosing/enclosed ellipse	
Stroma StdDev Diff. of enclosing/enclosed ellipse	
Stroma Mean Elliptic Fit	
Stroma StdDev Elliptic Fit	
Stroma Mean Length Pxl	
Stroma StdDev Length Pxl	
Stroma Mean Length/width	
Stroma StdDev Length/width	
Stroma Mean Main direction	
Stroma StdDev Main direction	
Stroma Mean Max.Diff.	30
Stroma StdDev Max.Diff.	
Stroma Mean Mean Channel 1	
Stroma StdDev Mean Channel 1	
Stroma Mean Mean Channel 2	
Stroma StdDev Mean Channel 2	
Stroma Mean Mean Channel 3	
Stroma StdDev Mean Channel 3	
Stroma Mean Radius of largest enclosed ellipse	
Stroma StdDev Radius of largest enclosed ellipse	
Stroma Mean Radius of smallest enclosing ellipse	
Stroma StdDev Radius of smallest enclosing ellipse	
Stroma Mean Rectangular Fit	40
Stroma StdDev Rectangular Fit	

【 0 3 4 6 】

【表 2 - 6】

フィーチャ	
Stroma Mean Shape index	
Stroma StdDev Shape index	
Stroma Mean Stddev Channel 1	
Stroma StdDev Stddev Channel 1	
Stroma Mean Stddev Channel 2	
Stroma StdDev Stddev Channel 2	
Stroma Mean Stddev Channel 3	
Stroma StdDev Stddev Channel 3	
Stroma Mean Width Pxl	10
Stroma StdDev Width Pxl	
Stroma Nuclei Mean Area Pxl	
Stroma Nuclei StdDev Area Pxl	
Stroma Nuclei Mean Asymmetry	
Stroma Nuclei StdDev Asymmetry	
Stroma Nuclei Mean Border index	
Stroma Nuclei StdDev Border index	
Stroma Nuclei Mean Border length Pxl	
Stroma Nuclei StdDev Border length Pxl	
Stroma Nuclei Mean Brightness	
Stroma Nuclei StdDev Brightness	
Stroma Nuclei Mean Compactness	20
Stroma Nuclei StdDev Compactness	
Stroma Nuclei Mean Density	
Stroma Nuclei StdDev Density	
Stroma Nuclei Mean Diff. of enclosing/enclosed ellipse	
Stroma Nuclei StdDev Diff. of enclosing/enclosed ellipse	
Stroma Nuclei Mean Elliptic Fit	
Stroma Nuclei StdDev Elliptic Fit	
Stroma Nuclei Mean Length Pxl	
Stroma Nuclei StdDev Length Pxl	
Stroma Nuclei Mean Length/width	
Stroma Nuclei StdDev Length/width	
Stroma Nuclei Mean Main direction	30
Stroma Nuclei StdDev Main direction	
Stroma Nuclei Mean Max.Diff.	
Stroma Nuclei StdDev Max.Diff.	
Stroma Nuclei Mean Mean Channel 1	
Stroma Nuclei StdDev Mean Channel 1	
Stroma Nuclei Mean Mean Channel 2	
Stroma Nuclei StdDev Mean Channel 2	
Stroma Nuclei Mean Mean Channel 3	
Stroma Nuclei StdDev Mean Channel 3	
Stroma Nuclei Mean Radius of largest enclosed ellipse	
Stroma Nuclei StdDev Radius of largest enclosed ellipse	
Stroma Nuclei Mean Radius of smallest enclosing ellipse	40
Stroma Nuclei StdDev Radius of smallest enclosing ellipse	

【 0 3 4 7 】

【表 2 - 7】

フィーチャ	
Stroma Nuclei Mean Rectangular Fit	
Stroma Nuclei StdDev Rectangular Fit	
Stroma Nuclei Mean Shape index	
Stroma Nuclei StdDev Shape index	
Stroma Nuclei Mean Stddev Channel 1	
Stroma Nuclei StdDev Stddev Channel 1	
Stroma Nuclei Mean Stddev Channel 2	
Stroma Nuclei StdDev Stddev Channel 2	
Stroma Nuclei Mean Stddev Channel 3	10
Stroma Nuclei StdDev Stddev Channel 3	
Stroma Nuclei Mean Width Pxl	
Stroma Nuclei StdDev Width Pxl	
Area of Artifact Pxl	
Area of Cytoplasm Pxl	
Area of Epithelial Nuclei Pxl	
Area of Lumen Pxl	
Area of Red Blood Cell Pxl	
Area of Stroma Pxl	
Area of Stroma Nuclei Pxl	
Number of objects of Artifact	
Number of objects of Cytoplasm	20
Number of objects of Epithelial Nuclei	
Number of objects of Lumen	
Number of objects of Red Blood Cell	
Number of objects of Stroma	
Number of objects of Stroma Nuclei	
Red Blood Cell Mean Area Pxl	
Red Blood Cell StdDev Area Pxl	
Red Blood Cell Mean Asymmetry	
Red Blood Cell StdDev Asymmetry	
Red Blood Cell Mean Border index	
Red Blood Cell StdDev Border index	
Red Blood Cell Mean Border length Pxl	30
Red Blood Cell StdDev Border length Pxl	
Red Blood Cell Mean Brightness	
Red Blood Cell StdDev Brightness	
Red Blood Cell Mean Compactness	
Red Blood Cell StdDev Compactness	
Red Blood Cell Mean Density	
Red Blood Cell StdDev Density	
Red Blood Cell Mean Diff. of enclosing/enclosed ellipse	
Red Blood Cell StdDev Diff. of enclosing/enclosed ellipse	
Red Blood Cell Mean Elliptic Fit	
Red Blood Cell StdDev Elliptic Fit	
Red Blood Cell Mean Length Pxl	40
Red Blood Cell StdDev Length Pxl	

【 0 3 4 8 】

【表 2 - 8】

フィーチャ	
Red Blood Cell Mean Length/width	
Red Blood Cell StdDev Length/width	
Red Blood Cell Mean Main direction	
Red Blood Cell StdDev Main direction	
Red Blood Cell Mean Max.Diff.	
Red Blood Cell StdDev Max.Diff.	
Red Blood Cell Mean Mean Channel 1	
Red Blood Cell StdDev Mean Channel 1	
Red Blood Cell Mean Mean Channel 2	10
Red Blood Cell StdDev Mean Channel 2	
Red Blood Cell Mean Mean Channel 3	
Red Blood Cell StdDev Mean Channel 3	
Red Blood Cell Mean Radius of largest enclosed ellipse	
Red Blood Cell StdDev Radius of largest enclosed ellipse	
Red Blood Cell Mean Radius of smallest enclosing ellipse	
Red Blood Cell StdDev Radius of smallest enclosing ellipse	
Red Blood Cell Mean Rectangular Fit	
Red Blood Cell StdDev Rectangular Fit	
Red Blood Cell Mean Shape index	
Red Blood Cell StdDev Shape index	
Red Blood Cell Mean Stddev Channel 1	20
Red Blood Cell StdDev Stddev Channel 1	
Red Blood Cell Mean Stddev Channel 2	
Red Blood Cell StdDev Stddev Channel 2	
Red Blood Cell Mean Stddev Channel 3	
Red Blood Cell StdDev Stddev Channel 3	
Red Blood Cell Mean Width Pxl	
Red Blood Cell StdDev Width Pxl	

【 0 3 4 9 】

【表 17】

表17. 形態計測的フィーチャ
Script v5.0(38フィーチャ)

フィーチャ名称	説明	
CytoplasmMeanMeanChannel40058	赤色チャンネルにおける平均細胞質強度値	
CytoplasmMeanMeanChannel50059	緑色チャンネルにおける平均細胞質強度値	
CytoplasmMeanMeanChannel60060	青色チャンネルにおける平均細胞質強度値	
CytoplasmMeanStddevChannel40066	赤色チャンネルにおける平均細胞質強度標準偏差	
CytoplasmMeanStddevChannel50067	緑色チャンネルにおける平均細胞質強度標準偏差	
CytoplasmMeanStddevChannel60068	青色チャンネルにおける平均細胞質強度標準偏差	
CytoplasmStddevMeanChannel40081	赤色チャンネルにおける細胞質強度の標準偏差	10
CytoplasmStddevMeanChannel50082	緑色チャンネルにおける細胞質強度の標準偏差	
CytoplasmStddevMeanChannel60083	青色チャンネルにおける細胞質強度の標準偏差	
EpithelialNucleiMeanAreaPx10101	上皮核領域の平均値(ピクセル)	
EpithelNucleiMeanMeanChanne40112	赤色チャンネルにおける平均上皮核強度	
EpithelNucleiMeanMeanChanne50113	緑色チャンネルにおける平均上皮核強度	
EpithelNucleiMeanMeanChanne60114	青色チャンネルにおける平均上皮核強度	
EpitheNucleiMeanStddevChann40120	赤色チャンネルにおける平均上皮核強度標準偏差	
EpitheNucleiMeanStddevChann50121	緑色チャンネルにおける平均上皮核強度標準偏差	
EpitheNucleiMeanStddevChann60122	青色チャンネルにおける平均上皮核強度標準偏差	
EpitheliaNucleiStddevAreaPx10124	上皮核領域の標準偏差	
EpitheNucleiStddevMeanChann40135	赤色チャンネルにおける上皮核強度の標準偏差	
EpitheNucleiStddevMeanChann50136	緑色チャンネルにおける上皮核強度の標準偏差	20
EpitheNucleiStddevMeanChann60137	青色チャンネルにおける上皮核強度の標準偏差	
StromaMeanMeanChannel40262	赤色チャンネルにおける平均間質強度	
StromaMeanMeanChannel50263	緑色チャンネルにおける平均間質強度	
StromaMeanMeanChannel60264	青色チャンネルにおける平均間質強度	
StromaMeanStddevChannel40270	赤色チャンネルにおける平均間質強度標準偏差	
StromaMeanStddevChannel50271	緑色チャンネルにおける平均間質強度標準偏差	
StromaMeanStddevChannel60272	青色チャンネルにおける平均間質強度標準偏差	
StromaStddevMeanChannel40331	赤色チャンネルにおける間質強度の標準偏差	
StromaStddevMeanChannel50332	緑色チャンネルにおける間質強度の標準偏差	
StromaStddevMeanChannel60333	青色チャンネルにおける間質強度の標準偏差	
AreaOfCytoplasmPx10345	細胞質の全領域(ピクセル)	
AreaOfEpithelialNucleiPx10350	上皮核の全領域(ピクセル)	
AreaOfLumenPx10357	管腔の全領域(ピクセル)	30
NumberOfObjectOfEpitheNuclei0364	上皮核オブジェクトの総数	
AreaCytodivTotTissueArea	組織領域に関する細胞質の相対領域(%)	
AreaEpitNucdivTotTissueArea	組織領域に関する上皮核の相対領域(%)	
AreaLumendivTotTissueArea	組織領域に関する管腔の相対領域(%)	
AreaStromadivTotTissueArea	組織領域に関する間質の相対領域(%)	
NumObjEpitNucdivTotNumberNuc	核の総数に関する上皮核の相対数	

【 0 3 5 0 】

【表 18 - 1】

表18. 形態計測的フィーチャ

フィーチャ名称	説明	
CytoplasmMeanMeanChannel40058	赤色フィルタを使用した細胞質強度平均値の平均	40
CytoplasmMeanMeanChannel50059	緑色フィルタを使用した細胞質強度平均値の平均	
CytoplasmMeanMeanChannel60060	青色フィルタを使用した細胞質強度平均値の平均	
CytoplasmMeanStddevChannel40066	赤色フィルタを使用した細胞質強度標準偏差の平均	
CytoplasmMeanStddevChannel50067	緑色フィルタを使用した細胞質強度標準偏差の平均	

【 0 3 5 1 】

【表 18 - 2】

フィーチャ名称	説明
CytoplasmMeanStddevChannel60068	青色フィルタを使用した細胞質強度標準偏差の平均
CytoplasmStddevMeanChannel40081	赤色フィルタを使用した平均細胞質強度の標準偏差
CytoplasmStddevMeanChannel50082	緑色フィルタを使用した平均細胞質強度の標準偏差
CytoplasmStddevMeanChannel60083	青色フィルタを使用した平均細胞質強度の標準偏差
EpithelNucleiMeanMeanChanne40112	赤色フィルタを使用した上皮核強度の平均
EpithelNucleiMeanMeanChanne50113	緑色フィルタを使用した上皮核強度の平均
EpithelNucleiMeanMeanChanne60114	青色フィルタを使用した上皮核強度の平均
EpitheNucleiMeanStddevChann40120	赤色フィルタを使用した上皮核強度標準偏差の平均
EpitheNucleiMeanStddevChann50121	緑色フィルタを使用した上皮核強度標準偏差の平均
EpitheNucleiMeanStddevChann60122	青色フィルタを使用した上皮核強度標準偏差の平均
EpitheliaNucleiStddevAreaPx10124	上皮核領域の標準偏差
EpitheNucleiStddevMeanChann40135	赤色フィルタを使用した平均上皮核強度の標準偏差
EpitheNucleiStddevMeanChann50136	緑色フィルタを使用した平均上皮核強度の標準偏差
EpitheNucleiStddevMeanChann60137	青色フィルタを使用した平均上皮核強度の標準偏差
StromaMeanMeanChannel40262	赤色フィルタを使用した間質強度の平均
StromaMeanMeanChannel50263	緑色フィルタを使用した間質強度の平均
StromaMeanMeanChannel60264	青色フィルタを使用した間質強度の平均
StromaMeanStddevChannel40270	赤色フィルタを使用した間質強度標準偏差の平均
StromaMeanStddevChannel50271	緑色フィルタを使用した間質強度標準偏差の平均
StromaMeanStddevChannel60272	青色フィルタを使用した間質強度標準偏差の平均
StromaStddevMeanChannel40331	赤色フィルタを使用した平均間質強度の標準偏差
StromaStddevMeanChannel50332	緑色フィルタを使用した平均間質強度の標準偏差
StromaStddevMeanChannel60333	青色フィルタを使用した間質強度の標準偏差
AreaCytodivTotTissueArea	組織領域に対する細胞質領域(%)
AreaEpitNucdivTotTissueArea	組織領域に対する上皮核領域(%)
AreaLumendivTotTissueArea	組織領域に対する管腔領域(%)
AreaRBCdivTotTissueArea	組織領域に対する赤血球領域(%)
AreaStromadivTotTissueArea	組織領域に対する間質領域(%)

10

20

【 0 3 5 2 】

【表 19 - 1】

表19. 実施例8 臨床情報

特徴	訓練	検証
N	342	340
術前PSA(ng/mL)		
平均	10.3	10.8
中央値	7.5	7.9
範囲	0.8-68.5	1.01- 00.0
リンパ節の関与		
陰性	339 (99.1%)	335 (98.5%)
陽性	3 (0.9%)	5(1.5%)
精囊の関与		
なし	320 (93.6%)	321 (94.4%)
あり	22 (6.4%)	19 (5.6%)

30

40

【 0 3 5 3 】

【表 19 - 2】

切除縁			
陰性	219 (64.0%)	240 (71.0%)	
陽性	123 (36.0%)	100 (29.0%)	
嚢外の関与			
なし	245 (71.6.0%)	254 (74.7%)	
あり	97 (28.4%)	86 (25.3%)	
優位生検グリーソン分類			
1	1 (0.3%)	2 (0.6%)	
2	17 (5.0%)	23 (6.8%)	
3	279 (81.6%)	273 (80.3%)	
4	45 (13.2%)	42 (12.4%)	
5	0 (0.0%)	0 (0.0%)	10
生検グリーソンスコア			
2	0 (0.0%)	1 (0.3%)	
3	1 (0.3%)	2 (0.6%)	
4	11 (3.2%)	8 (2.4%)	
5	23 (6.7%)	30 (8.8%)	
6	190 (55.6%)	187 (55.0%)	
7	101 (29.5%)	94 (27.7%)	
8	15 (4.4%)	13 (3.8%)	
9	1 (0.3%)	5 (1.5%)	
優位検体グリーソン分類			
2	8 (2.3%)	5 (1.5%)	
3	283 (82.8%)	276 (81.2%)	
4	49 (14.3%)	55 (16.2%)	
5	2 (0.6%)	4 (1.2%)	20
検体グリーソンスコア			
5	9 (2.6%)	15 (4.4%)	
6	116 (33.9%)	110 (32.4%)	
7	193 (56.4%)	187 (55.0%)	
8	18 (5.3%)	17 (5.0%)	
9	6 (1.8%)	10 (2.9%)	
10	0 (0.0%)	1 (0.3%)	

【 0 3 5 4 】

【表 20 - 1】

表20. 実施例9 評価および選択されたフィーチャ

フィーチャ	Aggregating 関数	一変量 CI*	CFモデルに おける加重
臨床的フィーチャ			
臨床的TNM期		0.32917	
術前PSA		0.30974	
優位生検グリーソン分類		0.3265	
生検グリーソンスコア		0.31104	
優位前立腺摘除グリーソン分類		0.26804	-16.4
前立腺摘除グリーソンの和		0.31311	
嚢外拡大		0.3358	
精嚢状態		0.37022	
リンパ節の関与		0.42033	-25.938
切除縁状態		0.37596	
形態計測的フィーチャ			
上皮核の相対領域		0.3778	

【 0 3 5 5 】

【表 2 0 - 2】

上皮核テクスチャフィーチャ1(正規化)		0.33503	
上皮核テクスチャフィーチャ2		0.39107	
上皮核テクスチャフィーチャ3		0.36808	
細胞質テクスチャフィーチャ1		0.72432	
細胞質テクスチャフィーチャ2		0.74648	20.728
細胞質テクスチャフィーチャ3		0.73593	
管腔の相対領域		0.67302	8.9663
管腔平均境界長		0.7646	17.217
単離上皮核密度		0.30897	
十分に定義された上皮核		0.34273	
単離上皮核領域(正規化)		0.3098	10
		0.42957	
分子(免疫蛍光)フィーチャ			
上皮核の相対領域	min	0.54152	
AR+上皮核の相対領域	mean	0.3778	
AMACR+上皮細胞の相対領域	max	0.45243	
AR+AMACR+/AR+上皮核の相対領域	max	0.50219	
AR+AMACR+/上皮核の相対領域	max	0.4362	
正規化和AR強度ピン3~9/総核領域	mean	0.33266	
正規化和AR強度ピン1~2/総核領域	max	0.60088	
正規化和AR強度ピン3~5/総核領域	mean	0.38799	20
正規化和AR強度ピン6~10/総核領域	min	0.45605	
AMACR-上皮核中の平均AR強度	mean	0.34475	-5.2075
AMACR+上皮核中の平均AR強度	mean	0.30174	
上皮核中の総AR強度	mean	0.32532	
上皮核中の正規化AR強度	mean	0.3406	
AMACR+上皮核中の総AR強度	max	0.42957	
AMACR-上皮核中の総AR強度	max	0.44616	
総AR強度×領域(正規化)	mean	0.34167	
総AR陽性強度×領域(正規化)	mean	0.33562	
上皮核内のAR+強度の全領域	mean	0.37898	30
* 訓練組中の臨床的な失敗の単一変数予測因子としての一致指数			

【 0 3 5 6 】

【表 2 1 - 1】

表21. 実施例9 患者の臨床的特徴

患者数	訓練		検証	
	頻度	比率(%)	頻度	比率(%)
<4	26	6.97	30	7.79
[4,10)	216	57.91	215	55.84
[10,20)	92	24.66	101	26.23
>20	39	10.46	39	10.13

【 0 3 5 7 】

【表 2 1 - 2】

臨床的TNM期	頻度	比率	頻度	比率
1	3	0.8	4	1.04
2	187	50.13	187	48.57
3	69	18.5	61	15.84
4	36	9.65	30	7.79
5	72	19.3	98	25.45
6	6	1.61	5	1.3
生検グリーソンスコア				
	頻度	比率	頻度	比率
2	0	0	1	0.26
3	1	0.27	2	0.52
4	11	2.95	9	2.34
5	23	6.17	31	8.05
6	202	54.16	205	53.25
7	113	30.29	108	28.05
8	22	5.9	23	5.97
9	1	0.27	6	1.56
前立腺摘除グリーソンの和				
	頻度	比率	頻度	比率
5	9	2.41	16	4.16
6	119	31.9	116	30.13
7	214	57.37	212	55.06
8	21	5.63	23	5.97
9	10	2.68	17	4.42
10	0	0	1	0.26
精嚢状態				
	頻度	比率	頻度	比率
0	342	91.69	349	90.65
1	31	8.31	36	9.35
切除縁状態				
	頻度	比率	頻度	比率
0	234	62.73	258	67.01
1	139	37.27	127	32.99
嚢外拡大				
	頻度	比率	頻度	比率
0	254	68.1	276	71.69
1	29	7.77	33	8.57
2	90	24.13	76	19.74
リンパ節の関与				
	頻度	比率	頻度	比率
0	364	97.59	373	96.88
1	9	2.41	12	3.12
打ち切り				
	頻度	比率	頻度	比率
0 (まだOF無し)	340	91.15	356	92.47
1 (CF)	33	8.85	29	7.53

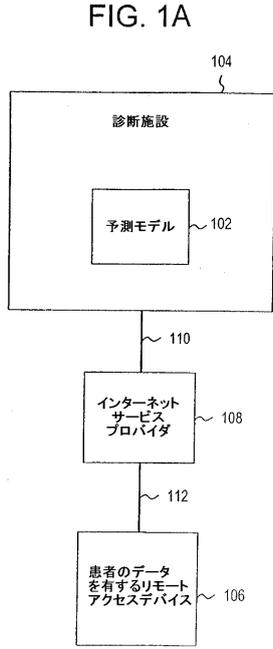
【 0 3 5 8 】

【表 2 2】

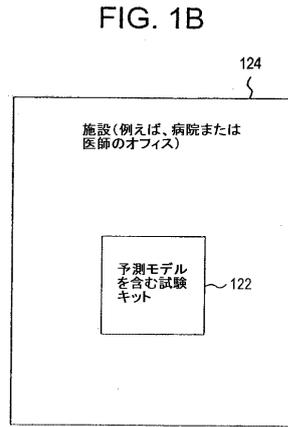
表22. 形態計測的フィーチャ-Script v6. 0(27フィーチャ)

フィーチャ名称	説明	
Cytoplasm Mean MeanChannel40058	赤色フィルタを使用した細胞質強度平均値の平均	
Cytoplasm Mean MeanChannel50059	緑色フィルタを使用した細胞質強度平均値の平均	
Cytoplasm Mean MeanChannel60060	青色フィルタを使用した細胞質強度平均値の平均	
CytoplasmMeanStddevChannel40066	赤色フィルタを使用した細胞質強度標準偏差の平均	
CytoplasmMeanStddevChannel50067	緑色フィルタを使用した細胞質強度標準偏差の平均	
CytoplasmMeanStddevChannel60068	青色フィルタを使用した細胞質強度標準偏差の平均	
CytoplasmStddevMeanChannel40081	赤色フィルタを使用した平均細胞質強度の標準偏差	10
CytoplasmStddevMeanChannel50082	緑色フィルタを使用した平均細胞質強度の標準偏差	
CytoplasmStddevMeanChannel60083	青色フィルタを使用した平均細胞質強度の標準偏差	
EpitheNucleiMeanStddevChann40142	赤色フィルタを使用した上皮核強度標準偏差の平均	
EpitheNucleiMeanStddevChann50143	緑色フィルタを使用した上皮核強度標準偏差の平均	
EpitheNucleiMeanStddevChann60144	青色フィルタを使用した上皮核強度標準偏差の平均	
EpitheNucleiStddevMeanChann40157	赤色フィルタを使用した平均上皮核強度の標準偏差	
EpitheNucleiStddevMeanChann50158	緑色フィルタを使用した平均上皮核強度の標準偏差	
EpitheNucleiStddevMeanChann60159	青色フィルタを使用した平均上皮核強度の標準偏差	
StromaMeanStddevChannel40310	赤色フィルタを使用した間質強度標準偏差の平均	
StromaMeanStddevChannel50311	緑色フィルタを使用した間質強度標準偏差の平均	20
StromaMeanStddevChannel60312	青色フィルタを使用した間質強度標準偏差の平均	
AreaCytopdivTotTissueArea	組織領域に対する細胞質領域(%)	
Area EpitNucdivTotTissueArea	組織領域に対する上皮核領域(%)	
AreaLumendivTotTissueArea	組織領域に対する管腔領域(%)	
AreaStromadivTotTissueArea	組織領域に対する間質領域(%)	
	核密度を、画像オブジェクトに中心があり、半径27ピクセルの円内の上皮核の相対領域と定義する。画像オブジェクトを、核密度値に基づいて10ピンに分類した。あらゆるピンについて、ピンを含むオブジェクトの全領域を計算した。	
AreaOfLowNucleiDensityNorm	フィーチャは、27ピクセル円内に40%までの上皮核を有する画像オブジェクトの相対領域を示す。	
AreaOfMedNucleiDensityNorm	フィーチャは、27ピクセル円内に40%超且つ60%未満の上皮核を有する画像オブジェクトの相対領域を示す。	30
AreaOfHighNucleiDensityNorm	フィーチャは、27ピクセル円内に60%超の上皮核を有する画像オブジェクトの相対領域を示す。	
DensityPercent_Bins3_4	半径27ピクセルの円内に20%超且つ40%未満の上皮核を有する画像オブジェクトの相対領域。	
DensityPercent_Bin1	半径27ピクセルの円内に10%未満の上皮核を有する画像オブジェクトの相対領域。	

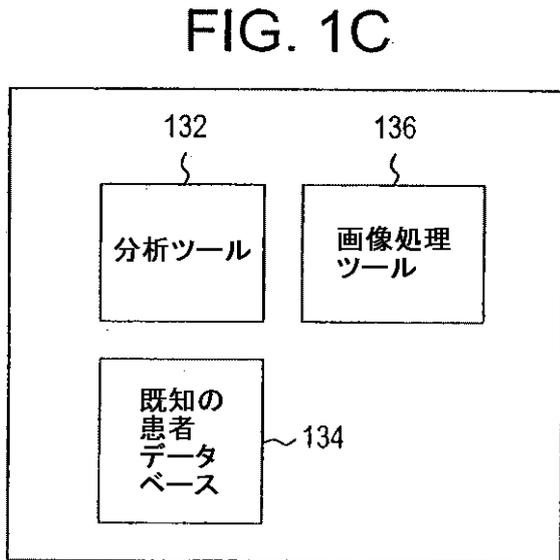
【図 1 A】



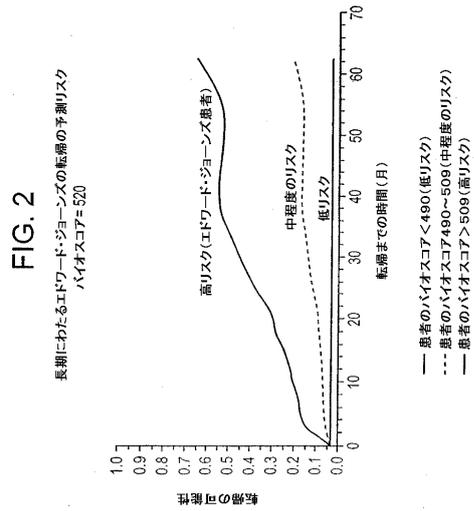
【図 1 B】



【図 1 C】

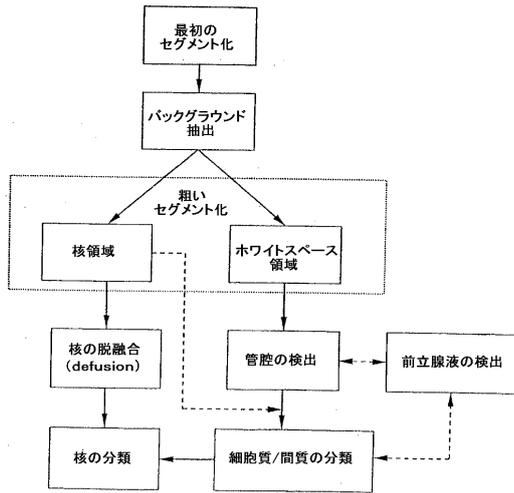


【図 2】



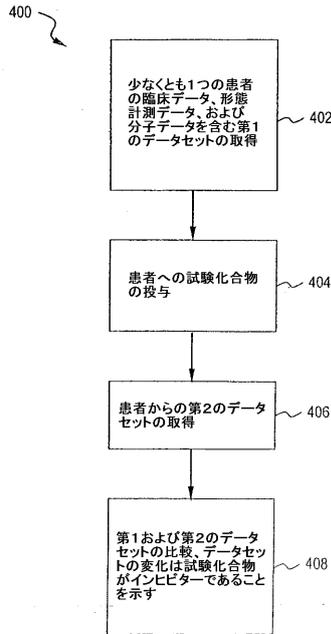
【 図 3 】

FIG. 3



【 図 4 】

FIG. 4



【 図 5 】

FIG. 5A

健康な前立腺組織の画像

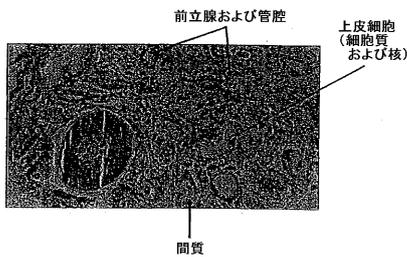
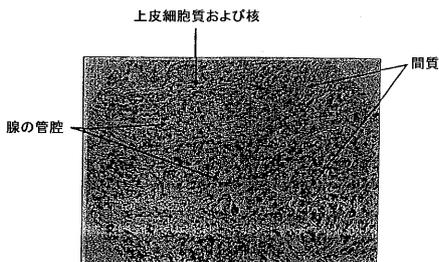


FIG. 5B

異常な前立腺組織の画像



【 図 6 】

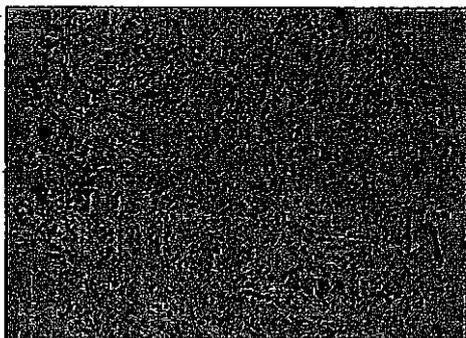
FIG. 6

PSA再発の予測のために選択されたフィーチャ(実施例2、研究1)

フィーチャ	説明
臨床的フィーチャ	生検グリーンスコア 人種 UICC期 倍数性の結果 DREの結果 リンパ節の関与 優位 (dominant) 生検グリーン分類 S期における倍数性の比率 術後グリーンスコア TNM期 優位術後グリーン分類 年齢 精囊の関与 術前PSA 倍数性画分の比率 切除線の関与 余外の関与
分子的フィーチャ	AR-腫瘍 AR-腺 CD34-腫瘍/PIN Ki67-腫瘍2 CD45-Fin 3 CD34-腫瘍/間質 Ki-67-腫瘍3 p27-腫瘍 C14-PIN CD34-腫瘍 PSA-腺 PSMA-PIN CD34-PIN/間質 CD45-腫瘍3
形態計測的フィーチャ	赤血球の最小長(ピクセル) 上皮核の最大稠密度 最小包囲(enclosure)の管腔最小半径 上皮核最小幅(ピクセル) 間質最大密度 管腔最小境界長(ピクセル) 上皮核最小標準偏差チャンネル2 最小包囲の上皮核最大半径 境界長の細胞質標準偏差(ピクセル) 領域のバックグラウンド標準偏差(ピクセル)

【 図 7 A 】

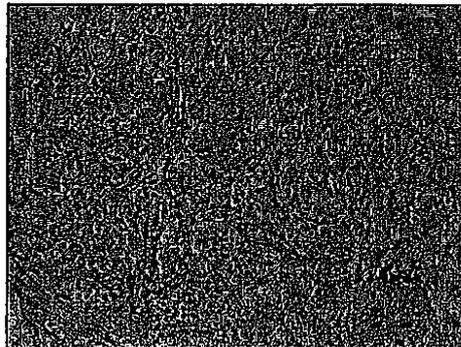
FIG. 7A



AR

【 図 7 B 】

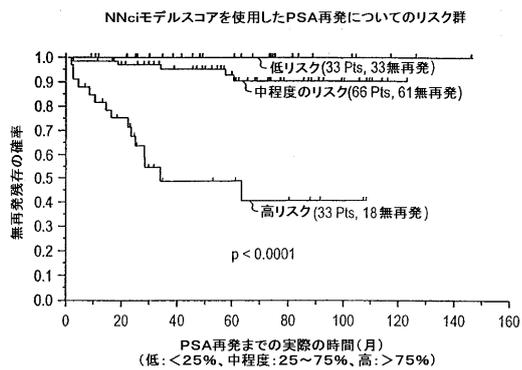
FIG. 7B



CD34

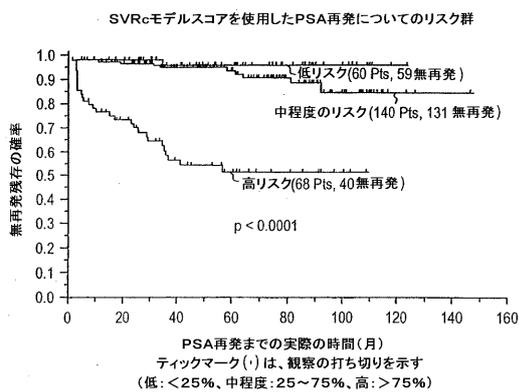
【 図 8 】

FIG. 8



【 図 10 】

FIG. 10



【 図 9 】

FIG. 9

PSA再発の予測のために選択されたフィーチャ(実施例2、研究2)

臨床的フィーチャ	TNM臨床期	0.7431
	切除線	0.7937
	リンパ節	0.8376
分子的フィーチャ	AR染色指標(腫瘍)	0.8528
形態計測的フィーチャ	EpithelialNucleiMinCompactne0215	0.8222
	StromaMaxStddevChannel30569	0.8483
	CytoplasmStddevMaxDiff0148	0.8569
	RedBloodCellMeanAreaPx10388	0.8596
	RedBloodCellStddevAreaPx10388	0.8621
	LumenMinAsymmetry0295	0.8635

【 図 1 1 】

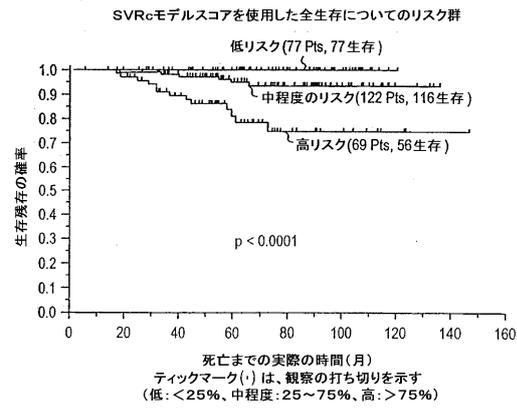
FIG. 11

全生存の予測のために選択されたフィーチャ(実施例2、研究3)

臨床的フィーチャ	tnm	0.7362
	年齢	0.7906
分子的フィーチャ	pspsi	0.7595
形態計測的フィーチャ	StromaMinMeanChannel10535	0.6804
	RedBloodCellMeanStddevChann30474	0.7475
	StromaMinMeanChannel20539	0.7722
	RedBloodCellMinMeanChannel20443	0.7772
	RedBloodCellStddevStddevChann20472	0.7809
	StromaMaxMaxDiff0529	0.7852
	EpithelNucleMeanBordeLengIPx10206	0.7888
	EpithelialNucleMeanAreaPx10194	0.7921
	EpithelNucleStddevElliptFit0228	0.7951
	RedBloodCellStddevStddevChann30476	0.7964
	RedBloodCellStddevElliptFit0420	0.7976

【 図 1 2 】

FIG. 12



【 図 1 3 】

FIG. 13

臨床的な失敗の予測のために選択されたフィーチャ(実施例3)

臨床的フィーチャ	<ul style="list-style-type: none"> 囊外抽出 精嚢浸潤 低位前立腺摘除グリーンソン分類 リンパ節浸潤
形態計測的フィーチャ	<ul style="list-style-type: none"> 総組織領域で除した細胞質領域 平均赤色チャネル (Red Channel) の細胞質標準偏差 総組織領域で除した管腔領域

【 図 1 4 】

FIG. 14

PSA再発の予測のために選択されたフィーチャ(実施例5)

臨床的フィーチャ	<ul style="list-style-type: none"> リンパ節 切除縁の関与 精嚢の関与 	<ul style="list-style-type: none"> -9.3742 -7.3159 -5.2103
分子的フィーチャ	AR(腫瘍)染色指標	-3.5404
形態計測的フィーチャ	<ul style="list-style-type: none"> AreaStromadivTotTissueArea AreaEpitNucdivTotTissueArea 	<ul style="list-style-type: none"> -3.4225 3.2975

【 図 1 6 】

FIG. 16

前立腺摘除後の臨床的な失敗の予測のために選択されたフィーチャ(実施例7)

臨床的フィーチャ	<ul style="list-style-type: none"> 生検グリーンソスコア リンパ節の関与 検体(前立腺摘除)のグリーンソスコア 	<ul style="list-style-type: none"> -18.08 -11.87 -6.36
分子的フィーチャ	AMACR上皮細胞内のAR強度	-6.20
形態計測的フィーチャ	<ul style="list-style-type: none"> CytoplasmMeanMeanChannel StromaMeanStddevChannel EpithelNucleStddevMeanChann 	<ul style="list-style-type: none"> 19.68 -10.06 -8.87

【 図 1 5 】

FIG. 15

PSA再発の予測のために選択されたフィーチャ(実施例6)

臨床的フィーチャ	<ul style="list-style-type: none"> リンパ節 切除縁の関与 生検グリーンソスコア 精嚢の関与 	<ul style="list-style-type: none"> -23.32 -11.73 -10.60 -6.40
分子的フィーチャ	AR(腫瘍)染色指標	-10.49
形態計測的フィーチャ	<ul style="list-style-type: none"> AreaStromadivTotTissueArea AreaEpitNucdivTotTissueArea StromaMeanStddevRedChann 	<ul style="list-style-type: none"> -16.15 11.54 -11.26

【 図 1 7 】

FIG. 17

PSA再発の予測のために選択されたフィーチャ(実施例8)

臨床的フィーチャ	<ul style="list-style-type: none"> 生検グリーンソスコア 精嚢浸潤 囊外抽出 術前PSA 低位前立腺摘除グリーンソン分類 	<ul style="list-style-type: none"> -25.60 -22.69 -4.37 -13.81 -10.93
分子的フィーチャ	AR+上皮核の相対領域	-10.64
形態計測的フィーチャ	<ul style="list-style-type: none"> 腫瘍上皮核のテクスチャ 腫瘍上皮細胞質のテクスチャ 	<ul style="list-style-type: none"> -19.21 18.19

【 図 1 9 】

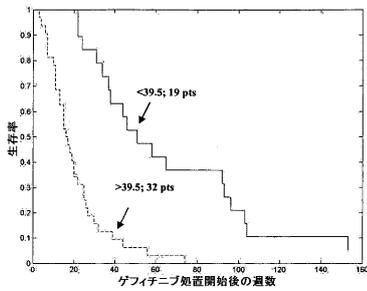
FIG. 19

前立腺摘除後の臨床的な失敗の予測のために選択されたフィーチャ(実施例9)

臨床的 フィーチャ	優位前立腺摘除グリーソン分類 リンパ節浸潤状態	-16.4 -25.938
分子的 フィーチャ	AR+/AMACR-上皮核における ARチャンネルの平均強度	-5.2075
形態計測的 フィーチャ	管腔の平均周囲長 グリーンチャンネルにおける細胞質内テクスチャの変動 管腔の相対面領域	17.217 20.728 8.9663

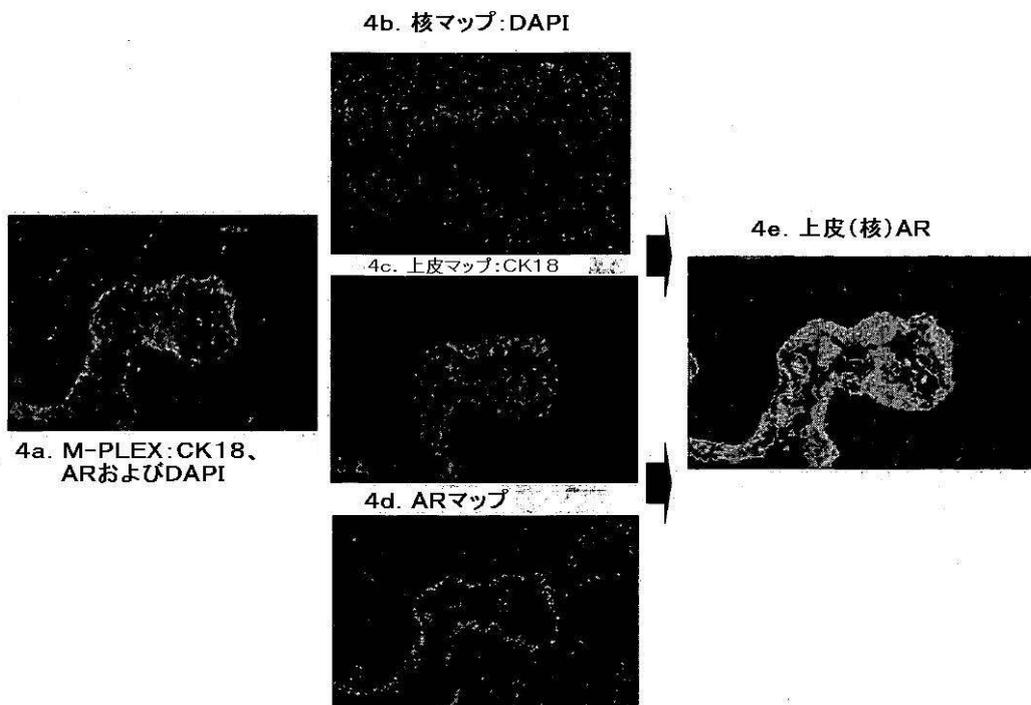
【 図 2 0 】

FIG. 20



【 図 1 8 】

FIG. 18



【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2008/004523

A. CLASSIFICATION OF SUBJECT MATTER INV. G06F19/00		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) G06F		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2005/262031 A1 (SAIDI OLIVIER [US] ET AL) 24 November 2005 (2005-11-24) paragraph [0008] - paragraph [0018] paragraph [0034] - paragraph [0259] figures 1,5B,6 -----	1-35
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. "&" document member of the same patent family		
Date of the actual completion of the international search 15 July 2008		Date of mailing of the international search report 22/07/2008
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer Rinelli, Pietro

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2008/004523

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2005262031	A1	NONE	24-11-2005

フロントページの続き

(31)優先権主張番号 61/010,598

(32)優先日 平成20年1月9日(2008.1.9)

(33)優先権主張国 米国(US)

(81)指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW

(72)発明者 テベロフスキ, ミハイル

アメリカ合衆国 ニューヨーク 10528, ハリソン, オズボーン ロード 190エー

(72)発明者 パーベル, デイビッド エー.

アメリカ合衆国 ニューヨーク 10021, ニューヨーク, 64ティーエイチ ストリート
412イー., アpartment 5ビー

(72)発明者 サイディ, オリビエ

アメリカ合衆国 コネチカット 06830, グリーンウィッチ, ウェスト パトナム アベ
ニュー 33, アpartment ナンバー2エイチ